

Reducing Non-uniqueness in Seismic Inverse Problems: New Observables in Seismology

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Geowissenschaften
der Ludwig-Maximilians-Universität München

vorgelegt von
Moritz Bernauer
aus Eggenfelden

am
19. Februar 2014

Erstgutachter: Prof. Dr. Heiner Igel

Zweitgutachter: Prof. Dr. Andreas Fichtner

Tag der mündlichen Prüfung: 10. Juli 2014

Summary

The scientific investigation of the solid Earth's complex processes, including their interactions with the oceans and the atmosphere, is an interdisciplinary field in which seismology has one key role. Major contributions of modern seismology are (1) the development of high-resolution tomographic images of the Earth's structure and (2) the investigation of earthquake source processes. In both disciplines the challenge lies in solving a seismic inverse problem, i.e. in obtaining information about physical parameters that are not directly observable.

Seismic inverse studies usually aim to find realistic models through the minimization of the misfit between observed and theoretically computed (synthetic) ground motions. In general, this approach depends on the numerical simulation of seismic waves propagating in a specified Earth model (forward problem) and the acquisition of illuminating data. While the former is routinely solved using spectral-element methods, many seismic inverse problems still suffer from the lack of information typically leading to ill-posed inverse problems with multiple solutions and trade-offs between the model parameters. Non-linearity in forward modeling and the non-convexity of misfit functions aggravate the inversion for structure and source.

This situation requires an efficient exploitation of the available data. However, a careful analysis of whether individual models can be considered a reasonable approximation of the true solution (deterministic approach) or if single models should be replaced with statistical distributions of model parameters (probabilistic or Bayesian approach) is inevitable.

Deterministic inversion attempts to find the model that provides the best explanation of the data, typically using iterative optimization techniques. To prevent the inversion process from being trapped in a meaningless local minimum an accurate initial low frequency model is indispensable. Regularization, e.g. in terms of smoothing or damping, is necessary to avoid artifacts from the mapping of high frequency information. However, regularization increases parameter trade-offs and is subjective to some degree, which means that resolution estimates tend to be biased.

Probabilistic (or Bayesian) inversions overcome the drawbacks of the deterministic approach by using a global model search that provides unbiased measures of resolution and trade-offs. Critical aspects are computational costs, the appropriate incorporation of prior knowledge and the difficulties in interpreting and processing the results.

This work studies both the deterministic and the probabilistic approach. Recent observations of rotational ground motions, that complement translational ground motion measurements from conventional seismometers, motivated the research. It is investigated if alternative seismic observables, including rotations and dynamic strain, have the potential to reduce non-uniqueness and parameter trade-offs in seismic inverse problems.

In the framework of deterministic full waveform inversion a novel approach to seismic tomography is applied for the first time to (synthetic) collocated measurements of translations, rotations and strain. The concept is based on the definition of new observables combining translation and

rotation, and translation and strain measurements, respectively. Studying the corresponding sensitivity kernels assesses the capability of the new observables to constrain various aspects of a three-dimensional Earth structure. These observables are generally sensitive only to small-scale near-receiver structures. It follows, for example, that knowledge of deeper Earth structure are not required in tomographic inversions for local structure based on the new observables.

Also in the context of deterministic full waveform inversion a new method for the design of seismic observables with focused sensitivity to a target model parameter class, e.g. density structure, is developed. This is achieved through the optimal linear combination of fundamental observables that can be any scalar measurement extracted from seismic recordings. A series of examples illustrate that the resulting optimal observables are able to minimize inter-parameter trade-offs that result from regularization in ill-posed multi-parameter inverse problems.

The inclusion of alternative and the design of optimal observables in seismic tomography also affect more general objectives in geoscience. The investigation of the history and the dynamics of tectonic plate motion benefits, for example, from the detailed knowledge of small-scale heterogeneities in the crust and the upper mantle. Optimal observables focusing on density help to independently constrain the Earth's temperature and composition and provide information on convective flow.

Moreover, the presented work analyzes for the first time if the inclusion of rotational ground motion measurements enables a more detailed description of earthquake source processes. The complexities of earthquake rupture suggest a probabilistic (or Bayesian) inversion approach. The results of the synthetic study indicate that the incorporation of rotational ground motion recordings can significantly reduce the non-uniqueness in finite source inversions, provided that measurement uncertainties are similar to or below the uncertainties of translational velocity recordings. If this condition is met, the joint processing of rotational and translational ground motion provides more detailed information about earthquake dynamics, including rheological fault properties and friction law parameters. Both are critical e.g. for the reliable assessment of seismic hazards.

Contents

1	Introduction	7
1.1	Living on a seismically active planet	7
1.2	Motivation of the thesis	8
1.2.1	Observing the complete ground motion	10
1.2.2	Inferring Earth's structure	10
1.2.3	Understanding earthquake rupture processes	11
1.3	Outline of the thesis	12
2	Measurements of translation, rotation and strain: new approaches to seismic processing and inversion	15
2.1	Introduction	15
2.2	The characteristics of a 12-component data set	17
2.3	New observables and their response to 3D Earth structure	18
2.4	Sensitivity Kernel Gallery	20
2.4.1	P wave kernels	20
2.4.2	S wave kernels	22
2.4.3	Love wave kernels	24
2.4.4	Rayleigh wave kernels	25
2.5	Perturbation Test	25
2.6	Discussion	28
2.7	Conclusions	29
3	Optimal observables for multi-parameter seismic tomography	31
3.1	Introduction	32
3.1.1	Multi-parameter seismic inversion	32
3.1.2	The origin of inter-parameter mappings in tomographic inversion	33
3.1.3	Outline	35
3.2	Theory	35
3.2.1	Constructing optimal observables	35

3.2.2	Parameter classes	38
3.3	Examples	38
3.3.1	Combining surface-wave traveltimes to enhance sensitivity to 3-D density variations	39
3.3.2	Sensitivity optimisation involving translational and rotational ground motion measurements	46
3.4	Discussion	49
3.4.1	Successfully designing optimal observables	50
3.4.2	The role of prior model covariance and parameter scaling	50
3.4.3	Iterative inversion for multiple parameter classes	51
3.4.4	Relation to the Backus-Gilbert method and experimental design	52
3.5	Conclusions	52
4	Reducing non-uniqueness in finite source inversion using rotational ground motions	55
4.1	Introduction	56
4.1.1	Non-uniqueness in finite source inversions	56
4.1.2	Rotational ground motions	56
4.1.3	Motivation and outline	57
4.2	Kinematic fault rupture modeling	58
4.2.1	Scenario and event geometry	58
4.2.2	The forward problem	58
4.2.3	Target model and synthetic data	60
4.3	Probabilistic finite source inversion	62
4.3.1	Bayesian inverse theory	62
4.3.2	Prior information and likelihood function	62
4.3.3	The Metropolis algorithm	63
4.3.4	Shannon's measure of information gain	64
4.4	Case studies with and without rotations	64
4.4.1	Scenario I: Inverting 3-C velocity observations at 20 stations	64
4.4.2	Scenario II: Inverting 6-C observations at 10 stations	67
4.4.3	Generalization: Random selection of stations	68
4.4.4	Requirements on the measurement accuracy of rotational ground motion sensors	68
4.5	Discussion	69
4.5.1	Measuring rotational ground motions	70
4.5.2	Physical origin of the information gain	71
4.5.3	Implications for earthquake physics and ground motion predictions	71
4.5.4	Probabilistic source inversion	71

4.6 Conclusions	72
Appendix	75

Chapter 1

Introduction

1.1 Living on a seismically active planet

"Before the discovery of Australia, people in the old world were convinced that all swans were white, an unassailable belief as it seemed completely confirmed by empirical evidence. The sighting of the first black swan might have been an interesting surprise for a few ornithologists (and others extremely concerned with the coloring of birds), but that is not where the significance of the story lies. It illustrates a severe limitation to our learning from observations or experience and the fragility of our knowledge. One single observation can invalidate a general statement derived from millennia of confirmatory sightings of millions of white swans. All you need is one single (and, I am told, quite ugly) black bird." (Taleb, 2007)

In the philosophical essay by Taleb (2007) the occurrence of a black swan describes a metaphor on hard-to-predict events with a high impact on the observer facilitated by ignorance. Paradoxically in earthquake history the appearance of black swans is the rule rather than the exception. Reviewing only the last decade we find a number of disastrous events. For example the M_w 9.1 Sumatra earthquake of 26 December 2004 generated a tsunami killing 300,000 people - many of them unaware of the phenomena preceding giant breaking waves. Deficient construction practices claimed a comparable number of fatalities and a damage of USD 7.8 billion in the M_w 7.0 Haiti earthquake of 12 January 2010. In the following year 20,000 casualties and an estimated damage of USD 300 billion in the M_w 9.0 Tohoku-Oki earthquake of 11 March 2011 demonstrated that not even highly-developed nations are safe from the violent forces suddenly released in portions of the Earth's crust. Moreover, intraplate events like the M_w 5.8 Virginia earthquake of 23 August 2011 frequently surprise and alert people that seismic threat is in general not restricted to plate-boundary locations. The scenarios indicate that the impact of earthquakes on human lives and infrastructure critically relies on the ignorance of the inhabitants of a seismically active planet with rapidly developing urban agglomerations (Bilham, 2009). However, mitigating the risk of earthquakes is a

complicated task. The complexity of earthquake triggering and the intrinsic variability of rupture processes severely hamper reliable hazard assessments (Stein *et al.*, 2012).

A better understanding of natural hazards such as earthquakes, tsunamis and volcanic eruptions requires a detailed knowledge of the processes in the solid Earth including their interaction with the oceans and the atmosphere. The scientific investigation of our planet's complex dynamics is crucial to raise awareness of natural phenomena and to build robustness to the corresponding perils. This is an interdisciplinary field in which seismology has one key role. Major contributions of modern seismology are

- the development of high-resolution tomographic images of the Earth's structure (seismic tomography) and
- the investigation of earthquake source processes.

In both disciplines the challenge lies in solving a seismic inverse problem, i.e. in obtaining information about physical parameters that are not directly observable.

1.2 Motivation of the thesis

Fundamental for inverting both Earth's structure and earthquake rupture models is the misfit quantification between observed and theoretically computed (synthetic) ground motions. Unfortunately in realistic Earth models the elastic wave equation has no analytical solution. Hence an essential task in seismology is the numerical simulation of seismic waves propagating in a specified Earth model (forward problem).

The forward problem is largely solved in seismology. For an accurate approximation of the complete wave field including body, surface and scattered waves in complex three-dimensional (3-D) heterogeneous media the spectral-element method (e.g. Komatitsch & Tromp, 1999; Komatitsch *et al.*, 2005; Chaljub *et al.*, 2007; Fichtner & Igel, 2008) has been established because of its computational efficiency. The method rests upon a weak formulation of the wave equation where the displacement field is approximated in terms of Lagrange polynomials with Gauss-Lobatto-Legendre collocation points (Fig. 1.1). This approach implicitly accounts for the free surface boundary condition and results in a diagonal mass matrix, which substantially reduces the complexity of the algorithm. Considering a spherically symmetric Earth model seismic wave propagation can be simulated even globally for frequencies up to 1 Hz with spectral-element solvers (e.g. Nissen-Meyer *et al.*, 2007). For 3-D heterogeneous models the computational domain is often subdivided into hexahedral elements and allows the meshing of surface topography and layer boundaries in the Earth model.

The strategies developed in this project do not rely on a specific forward solver. However, combining accuracy and flexibility the spectral-element method on a regular hexahedral grid for poten-

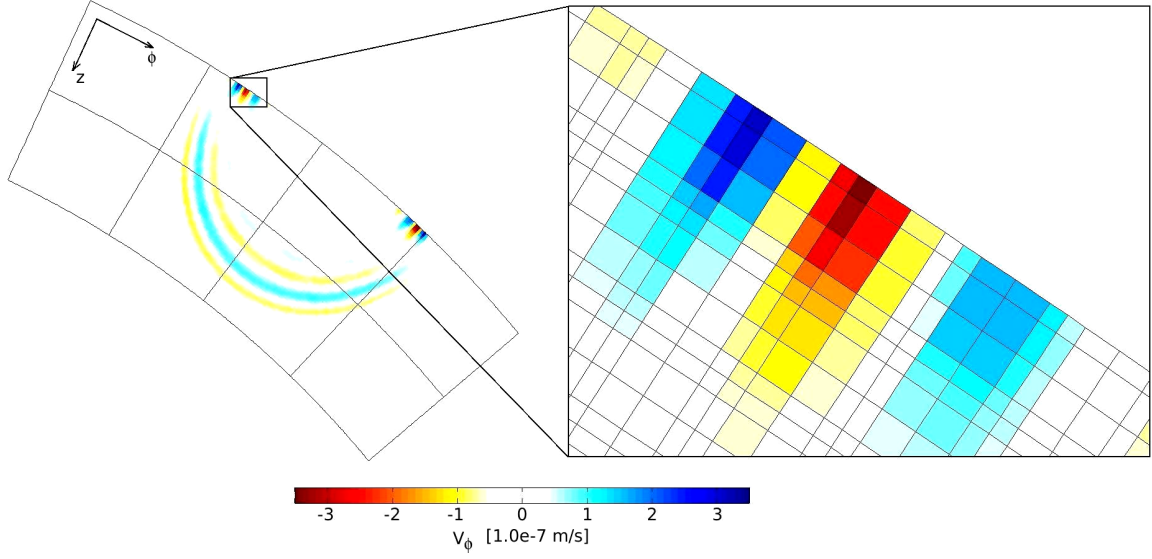


Figure 1.1: Numerical simulation of seismic wave propagation with a spectral-element solver on a regular hexahedral grid. The vertical slice through a spherical Earth section illustrates a horizontal component of the velocity field. The spatial discretization is based on Gauss-Lobato-Legendre collocation points (zoom-in).

tially complex 3-D models has proved most appropriate for seismic wave field simulations in the presented test cases.

Having solved the forward problem the solution of a seismic inverse problem essentially depends on the available amount of data comparable with synthetics. Especially on continental and global scales seismic tomography often suffers from the uneven distribution of sources and receivers. Also the investigation of the rupture processes of large earthquakes is aggravated by the sparse observation of strong ground motions. Typically, in both cases the lack of data leads to an ill-posed inverse problem with multiple solutions and trade-offs between the model parameters. This requires a careful analysis of whether individual models can be considered a reasonable approximation of the true solution (deterministic approach) or if single models should be replaced with statistical distributions of model parameters (probabilistic or Bayesian approach).

To reduce the non-uniqueness in ill-posed inverse problems it is essential to increase the quantity and the quality of observations. Ambitious instrumentation projects like the EarthScope USArray in the United States (www.usarray.org) or the deployment of ocean-bottom seismometers illuminating thus far poorly sampled regions in the Indian Ocean by the RHUM-RUM project (www.rhum-rum.net) are important contributions in that context. However, conventional seismometers usually record three components (3-C) of translational ground motions (or their time derivatives) in North-South, East-West and vertical direction. In classical elasticity theory with infinitesimal deformations - adopted in most seismic applications - the complete description of ground motion depends on twelve components (12-C) (e.g. Aki & Richards, 2002). This potential diversity of measurements

complementing seismic data sets has stimulated the presented work.

1.2.1 Observing the complete ground motion

In addition to 3-C translational ground motions seismic tremor induces rotational ground motions around three orthogonal axes (3-C) and dynamic strain (6-C). While dynamic strain has been recorded for decades (Benioff *et al.*, 1951; Gombert & Agnew, 1996; Agnew & Wyatt, 2003) the observation of ground rotations from teleseismic earthquakes (Igel *et al.*, 2005, 2007) and even Earth's free oscillations (Igel *et al.*, 2011) has only recently become feasible with modern ring-laser gyroscopes (Schreiber *et al.*, 2006, 2009). Successful observatory measurements in the range of nanoradians (10^{-9}) per second and smaller initiated the development of field-deployable rotation sensors for seismological and engineering applications (e.g. Bernauer *et al.*, 2012; Schreiber *et al.*, 2009; Velikoseltsev *et al.*, 2012).

The incorporation of this new observable into seismic studies has evolved into an active research field. The amplitude ratio of collocated measurements of rotational and translational ground motions for instance revealed its potential in numerous applications. Considering jointly processed rotations and translations from teleseismic records leads to attractive observables for seismic tomography with local sensitivity and without the use of traveltimes (Fichtner & Igel, 2009; Bernauer *et al.*, 2009). The same approach allows the back azimuth of seismic events to be determined (Igel *et al.*, 2007) and can be used to estimate the direction of Love waves in ocean-generated noise (Hadziioannou *et al.*, 2012). The latter is for example relevant for passive imaging methods promising high resolution information about the crust and uppermost mantle (Shapiro *et al.*, 2005; Ritzwoller, 2009).

Motivated by these discoveries a major research topic of the thesis is to further investigate if alternative seismic observables (rotations and strain) complementing conventional seismic data (translations) are useful for inverting Earth's structure and earthquake source characteristics.

1.2.2 Inferring Earth's structure

Besides the efforts to collect more data a substantial progress in modern seismology is the inversion of structural parameters from complete seismograms. Based on the continuously increasing computational power, newly developed tomographic methods, for example combining spectral-element and adjoint techniques, allow the analysis of full waveform misfits between observed and synthetic seismograms (e.g. Tape *et al.*, 2009; Zhu *et al.*, 2012; Fichtner *et al.*, 2010, 2013; Rickers *et al.*, 2013).

Nevertheless, minimizing the misfit between data and synthetics is a non-linear optimization problem depending on non-convex misfit functions. The challenge in solving such an optimization problem is to find the global minimum instead of being trapped in a local one. In the deterministic approach this problem is commonly solved iteratively by applying gradient-type optimization

schemes, i.e. at several iteration steps an initial model is updated according to the gradient of the misfit function computed with respect to the model parameters. In practice the gradient is composed of finite-frequency (also named sensitivity or Fréchet) kernels that can be calculated efficiently with the adjoint method (Tromp *et al.*, 2005; Fichtner *et al.*, 2006a,b; Sieminski *et al.*, 2007b,a). To avoid the convergence to local minima an accurate initial low frequency model is necessary. Then successively higher frequencies may be included into the inversion process to recover also small-scale structures. For instance, these are important to constrain the details of the crust and the upper mantle that are supposed to provide insights into the history and the dynamics of tectonic plate motion (Fichtner *et al.*, 2013) and also concern global tomography models (Ferreira *et al.*, 2010).

However, in ill-posed inverse problems a finite data set naturally limits the resolution of the model parameters (e.g. Backus & Gilbert, 1967). To prevent full waveform inversion techniques from mapping high frequency information into artificial Earth structure regularization, for example in terms of smoothing or damping, is necessary. The drawback is that regularization increases parameter trade-offs and is to some degree subjective, meaning that resolution estimates tend to be biased (Trampert & Fichtner, 2013).

It is therefore crucial to advance the development of optimal analysis tools that increase the resolution power in tomographic models and thus contribute to a comprehensive characterization of the Earth's internal state. In general, this is a multi-parameter inverse problem depending on the combined knowledge of a broad spectrum of physical properties. Joint information on elastic parameters and density, for instance, is needed to independently constrain temperature and composition (e.g. Trampert *et al.*, 2004; Mosca *et al.*, 2012) and lateral variations in density provide information on convective flow (e.g. Ishii & Tromp, 2001; Mosca *et al.*, 2012). Hence a key aspect of this study is the design of optimal observables that focus on target parameter classes in multi-parameter inverse problems.

1.2.3 Understanding earthquake rupture processes

In tomographic studies the release of seismic energy is usually idealized in terms of point sources described by the corresponding hypocenter coordinates and the moment tensor solution. However, a better understanding of seismic hazard and earthquake dynamics requires a more detailed description of the kinematics of extended earthquake sources. Thus the evolution of rupture processes in space and time is often modeled as a dislocation (slip) on planar fault surfaces specified by the rupture velocity and the rise time, i.e. the duration of the slip (e.g. Spudich & Archuleta, 1987).

Neglecting e.g. dynamic source models with physically self-consistent strain and stress fields, inelastic deformations of rocks, thermal expansion or frictional melting on the rupture plane during the earthquake, this approach is already a considerable simplification to the true situation. Nevertheless, it defines a severely ill-posed inverse problem. Non-linearity in the forward model, noisy data with insufficient coverage, and trade-offs between model parameters, result in a well-documented

non-uniqueness of finite source models (e.g. Olson & Anderson, 1988; Beresnev, 2003; Ide *et al.*, 2005; Mai *et al.*, 2007) and render deterministic solutions largely unfeasible.

The complexities of earthquake rupture processes combined with the sparse data situation rather suggest concentrating on a comprehensive management of the lack of knowledge, which is provided by a probabilistic (or Bayesian) approach. Though the proper inclusion of prior experience is of critical importance in probabilistic inversion (Tarantola, 2006; Efron, 2013) Bayesian source inversion overcomes the drawbacks of the deterministic approach using a global model search and provides unbiased measures of resolution and trade-offs.

A more detailed analysis of earthquake dynamics including the inference of rheological fault properties and friction law parameters is crucial e.g. for the reliable assessment of seismic hazard. So this study aims to investigate in a probabilistic framework whether jointly processed rotational and translational ground motion measurement are able to reduce the non-uniqueness and trade-offs in kinematic source inversions.

1.3 Outline of the thesis

With a strong focus on the inclusion of alternative seismic observables this work investigates how to reduce the non-uniqueness and the trade-offs in deterministic and probabilistic methods for inferring Earth's structure and understanding earthquake rupture processes.

Chapter 2 applies for the first time a novel approach in seismic tomography to collocated 12-C measurements, i.e. translations, rotations and strain. The concept is based on the apparent S and P velocities, defined as the ratios of displacement velocity and rotation amplitude, and displacement velocity and divergence amplitude, respectively. Studying their corresponding finite-frequency kernels, computed with a combination of spectral-element simulations and adjoint techniques, assess the capability of these new observables to constrain various aspects of 3-D Earth structure. The principal conclusion is that both the apparent S and P velocities are generally sensitive only to small-scale near-receiver structures, irrespective of the type of seismic wave considered. Hence knowledge of deeper Earth structure is not required in tomographic inversions for local structures based on the new observables. Moreover, a synthetic finite-perturbation test confirms the ability of the apparent S and P velocities to directly detect both the location and the sign of shallow lateral velocity variations.

Chapter 3 introduces a method for the design of seismic observables with maximum sensitivity to a target model parameter class, and minimum sensitivity to all remaining parameter classes. This is achieved through the optimal linear combination of fundamental observables that can be any scalar measurement extracted from seismic recordings, including frequency-dependent traveltimes and amplitudes of arbitrary seismic phases. The optimal weights of the fundamental observables

are determined with an efficient global search algorithm. A series of examples illustrate that the resulting optimal observables are able to minimize inter-parameter trade-offs that result from regularization in ill-posed multi-parameter inverse problems. To the benefit of reproducibility and clarity, the examples include a small number of well-understood fundamental observables, few parameter classes and a radially symmetric Earth model. The method itself, however, does not impose such restrictions, and it can easily be applied to large numbers of fundamental observables and parameter classes, as well as to 3-D heterogeneous Earth models.

Chapter 4 assesses the potential of rotational ground motion measurements to reduce non-uniqueness in kinematic source inversions. Based on synthetic Bayesian finite source inversions the results indicate that the incorporation of rotational ground motion recordings can significantly reduce non-uniqueness in finite source inversions, provided that measurement uncertainties are similar to or below the uncertainties of translational velocity recordings. When this condition is met, rupture velocity and rise time have the highest potential to benefit from rotation data. The trade-offs between both parameters can be strongly reduced, and the information gain significantly increased. It follows that rotation data can improve secondary inferences that depend on rupture velocity and rise time, such as radiation directivity and frictional properties of the fault.

Chapters 2, 3 and 4 of this thesis correspond to papers published in or submitted to peer-reviewed journals and can be read independently. In order of their appearance the relevant papers are:

- Bernauer, M., Fichtner, A. and Igel, H., 2012. Measurements of translation, rotation and strain: new approaches to seismic processing and inversion. *J. Seismol.*, **16**, 669-681
- Bernauer, M., Fichtner, A. and Igel, H., 12/2013. Optimal observables for multi-parameter seismic tomography. submitted to *Geophys. J. Int.*
- Bernauer, M., Fichtner, A. and Igel, H., 02/2014. Reducing non-uniqueness in finite source inversion using rotational ground motions. submitted to *J. Geophys. Res.*

Chapter 2

Measurements of translation, rotation and strain: new approaches to seismic processing and inversion

The content of this chapter is published as: Bernauer, M., Fichtner, A. and Igel, H., 2012. Measurements of translation, rotation and strain: New approaches to seismic processing and inversion. *J. Seismol.*, **16**, 669-681.

Summary

We propose a novel approach to seismic tomography based on the joint processing of translation, strain and rotation measurements. Our concept is based on the apparent S and P velocities, defined as the ratios of displacement velocity and rotation amplitude, and displacement velocity and divergence amplitude, respectively. To assess the capability of these new observables to constrain various aspects of 3D Earth structure, we study their corresponding finite-frequency kernels, computed with a combination of spectral-element simulations and adjoint techniques. The principal conclusion is that both the apparent S and P velocities are generally sensitive only to small-scale near-receiver structure, irrespective of the type of seismic wave considered. It follows that knowledge of deeper Earth structure would not be required in tomographic inversions for local structure based on the new observables. In a synthetic finite-perturbation test we confirm the ability of the apparent S and P velocities to directly detect both the location and the sign of shallow lateral velocity variations.

2.1 Introduction

Thanks to recent technological developments, seismically induced rotation and strain are emerging as new observables that complement the traditional translation recordings. Dynamic strain can be

recorded by long-base laser strainmeters (e.g. Agnew & Wyatt, 2003), and ring lasers are used for high-precision measurements of rotational ground motions (e.g. Schreiber *et al.*, 2009). Direct and array-derived rotation measurements were compared for ring laser systems (e.g. Suryanto *et al.*, 2006) and recordings of seismically induced strain have been analysed together with seismometer data and theoretical computations by Gombert & Agnew (1996). For dynamic strain measurements also entire station networks as for example the EarthScope borehole strainmeter array exist. The testing of portable rotation sensors has started only recently (e.g. Brokešová & Málek, 2010; Nigbor *et al.*, 2009; Liu *et al.*, 2009; Wassermann *et al.*, 2009).

While e.g. Mikumo & Aki (1964), Sacks *et al.* (1976) and Blum *et al.* (2010) showed how to derive local phase velocities from acceleration measurements in conjunction with dynamic strain observations also the newly developed rotation sensors have opened remarkable perspectives in many branches of seismological research: Observations of near-field rotational ground motions induced by swarm quakes (Takeo, 2009) and fault ruptures (Wu *et al.*, 2009) are likely to contribute to our understanding of earthquake source processes. As suggested by Pillet *et al.* (2009), rotation sensors may be used to improve the signal-to-noise ratio of ocean-bottom seismometers.

Also, various methods to infer Earth structure from rotation measurements have recently been proposed: Collocated measurements of translations and rotations were used to estimate local phase velocities (Igel *et al.*, 2005, 2007; Cochard *et al.*, 2006) or to identify the low seismic velocities of sedimentary basins (Wang *et al.*, 2009; Stupazzini *et al.*, 2009). Pham *et al.* (2009) extracted information about crustal scattering from rotational signals in the coda of P-waves. Ferreira & Igel (2009) used full ray-theory modelling to demonstrate a clearly observable effect of near-receiver heterogeneities on rotational motions of Love waves. Following these first successful applications based on single rotational ground motion recordings, the next steps consist in (1) the installation of rotation sensor networks, and (2) the incorporation of strain measurements in order to complete the set of seismic observables.

In anticipation of these developments, this paper explores the potentials of future rotation and strain sensor networks in the context of seismic tomography. For this we investigate an approach to seismic tomography that is based on the joint processing of translation (u_i), strain ($e_{ij} := \frac{1}{2}(\partial_j u_i + \partial_i u_j)$) and rotation ($\omega_i := \frac{1}{2}\epsilon_{ijk}\partial_j u_k$) measurements. Following Fichtner & Igel (2009) and Bernauer *et al.* (2009), we consider the apparent S and P velocities, defined as the ratios of displacement velocity and rotation amplitude, and displacement velocity and divergence amplitude, respectively. Using adjoint techniques and a spectral-element solver of the seismic wave equation (Fichtner *et al.*, 2009; Fichtner, 2010), we compute sensitivity kernels for the apparent P and S velocities. For a 1D-model the sensitivity of the apparent S velocity of S and surface waves is concentrated in the vicinity of the receiver (Fichtner & Igel, 2009). In this study we extend the work of Fichtner & Igel (2009) to P waves and the apparent P velocity. Furthermore, we combine the kernels with a finite-perturbation test indicating that aspects of 3D Earth structure are particularly well constrained by the newly defined observables.

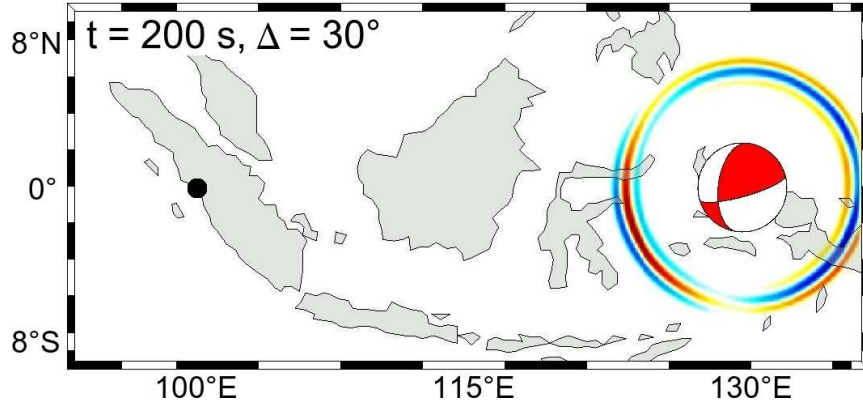


Figure 2.1: Source mechanism and snapshot of the vertical-component wave field, 200 s after source initiation. The source is located at 400 km depth, 130° E longitude and 0° latitude in eastern Indonesia. The dominant period is 20 s. The receiver is marked by the black dot (100° E longitude, 0° latitude, 0 km depth)

2.2 The characteristics of a 12-component data set

We begin our developments with the simulation of a 12-component data set that illustrates several characteristic properties of rotation and strain measurements. For the computation of both synthetic seismograms and sensitivity kernels we use a spectral-element solver of the seismic wave equation (Fichtner *et al.*, 2009; Fichtner, 2010). As Earth model we use the isotropic version of PREM without visco-elastic dissipation (Dziewoński & Anderson, 1981). The seismic wave field with a dominant period of 20 s is excited by a point source at 400 km depth beneath eastern Indonesia, and recorded at 30° epicentral distance, as illustrated in Fig.2.1. The explicit source parameters can be found in Appendix A.

The three-component displacement velocity, three-component rotation and six-component strain corresponding to the previously described setup are displayed in Fig. 2.2. The symbols θ , ϕ and z denote colatitude, longitude and depth, respectively. The arrival of the direct P wave around 330 s is visible in the radial and vertical components of the displacement velocity v_ϕ and v_z , as well as in the $\phi\phi$ and zz -components of the strain tensor. P-to-S conversions at the free surface are responsible for the P wave signal in the transverse rotation ω_θ . The clearest S wave arrival at ~ 600 s is contained in the transverse velocity v_θ , the vertical rotation ω_z , and the $\phi\phi$ and $\theta\phi$ strain components. Love and Rayleigh waves are present roughly from 700 s to 800 s.

While the waveforms of the non-zero strain and rotation components strongly depend on the source-receiver geometry, the vanishing components are of a more general nature. In particular, the radial-component rotation - ω_ϕ in our case - is always expected to be zero in a layered medium (Cochard *et al.*, 2006). Furthermore, the free surface boundary condition forces the strain components $e_{\phi z}$ and $e_{\theta z}$ to zero, which leads us to focus on the diagonal strain components in our subsequent analysis.

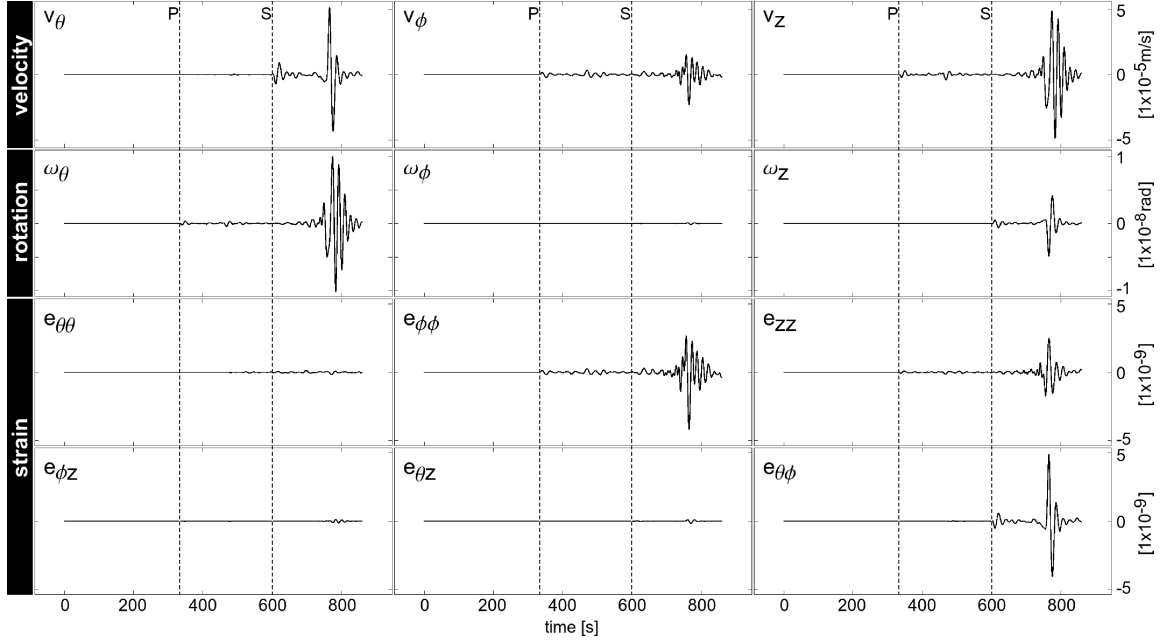


Figure 2.2: Synthetic 12-component data set: velocity, rotation and strain are simulated at one single seismic station on Sumatra. The dominant period is 20 s. P and S phases are marked by dashed lines

2.3 New observables and their response to 3D Earth structure

While rotation and strain measurements are interesting already by themselves, we wish to go one step further and define new observables from combinations of velocity, strain and rotation. This is motivated by a simple plane wave analysis: Assuming a plane S wave $\mathbf{u}(\mathbf{x}, t)$ in a homogeneous and isotropic full space, the S velocity β can be expressed as the ratio of velocity and rotation amplitude:

$$\beta = \frac{1}{2} \frac{|\mathbf{v}|}{|\boldsymbol{\omega}|}, \quad (2.1)$$

with the time derivative of the displacement field $\mathbf{v} = \dot{\mathbf{u}}$. Similarly, for a plane P wave, the P velocity α is given by

$$\alpha = \frac{|\mathbf{v}|}{|\text{tr } \mathbf{e}|}. \quad (2.2)$$

Due to our restrictive assumptions, equations 2.1 and 2.2 are of little practical relevance in heterogeneous media. Nevertheless, a slight generalisation promises to yield rather direct information on the Earth's S and P velocity structure: Inspired by equation 2.1 we define the apparent S velocity

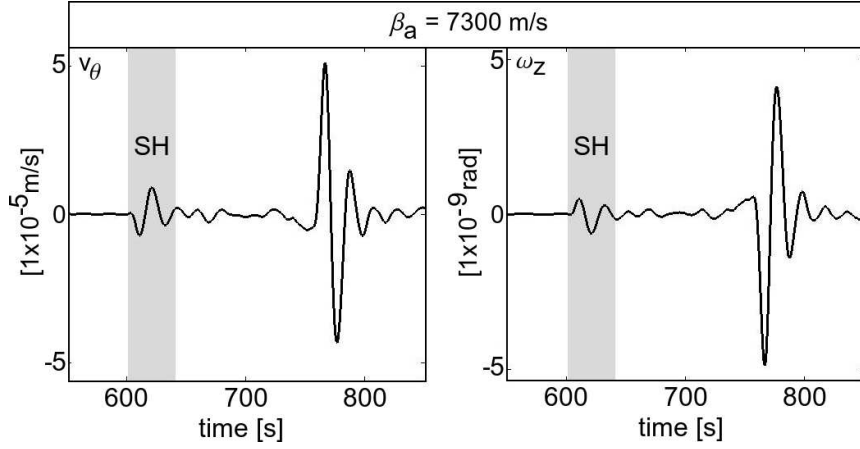


Figure 2.3: β_a measurement representing the SH wave phases of the v_θ and ω_z seismograms in figure 2.2. According to equation 2.3 β_a is calculated via $\frac{1}{2} \sqrt{\int v_\theta(\mathbf{x}^r, t)^2 dt} / \sqrt{\int \omega_z(\mathbf{x}^r, t)^2 dt}$, while v_θ and ω_z are restricted to the SH wave windows (grey column)

β_a measured at receiver position \mathbf{x}^r as (Fichtner & Igel, 2009)

$$\beta_a(\mathbf{x}^r) = \frac{1}{2} \frac{v(\mathbf{x}^r)}{w(\mathbf{x}^r)}, \quad (2.3)$$

with the velocity amplitude $v(\mathbf{x}^r) = \sqrt{\int \mathbf{v}^2(\mathbf{x}^r, t) dt}$,

and the rotation amplitude $w(\mathbf{x}^r) = \sqrt{\int \omega^2(\mathbf{x}^r, t) dt}$.

In analogy to equation 2.3 we define the apparent P velocity α_a at the receiver position \mathbf{x}^r as

$$\alpha_a(\mathbf{x}^r) = \frac{v(\mathbf{x}^r)}{s(\mathbf{x}^r)}, \quad (2.4)$$

where the symbol s denotes the divergence amplitude $s(\mathbf{x}^r) = \sqrt{\int [\text{tr} \mathbf{e}(\mathbf{x}^r, t)]^2 dt}$. Both definitions, 2.3 and 2.4, are applicable either to complete seismograms or to isolated waveforms, as discussed in section 2.4. An exemplary measurement of β_a corresponding to the SH wave phases of the v_θ and ω_z seismograms in figure 2.2 is shown in Figure 2.3. The apparent S velocity β_a is equal to the S velocity β in the case of a plane S wave in a homogeneous, unbounded and isotropic medium. A similar result holds for the apparent P velocity α_a .

It is at this point important to keep in mind that the apparent P and S velocities α_a and β_a are measurements derived from various seismograms. In contrast, the P and S velocities α and β are material parameters, the 3D variations of which are the target of a tomographic inversion.

Our primary interest is in the response of the previously defined measurements to variations in 3D Earth structure. For this, we consider a generic measurement χ that represents, for instance, the velocity amplitude $v(\mathbf{x}^r)$, the apparent S velocity $\beta_a(\mathbf{x}^r)$, or the apparent P velocity $\alpha_a(\mathbf{x}^r)$. A change in the observable, $\delta\chi$, that results from a model perturbation δm is given, correct to first

order, by

$$\delta\chi = \int K_m(\chi, \mathbf{x}) \delta m(\mathbf{x}) d^3\mathbf{x}, \quad (2.5)$$

where m can be the S velocity β , the P velocity α or any other Earth model parameter. The sensitivity or Fréchet kernel $K_m(\chi, \mathbf{x})$ describes how the observable χ is affected by model parameter changes δm at position \mathbf{x} in the Earth. In the interest of a succinct notation we omit the spatial dependence of the kernels from hereon.

In the following we explore the properties of sensitivity kernels for various rotation- and strain-related observables, including the apparent S and P velocities. This allows us to identify the observable's capability to constrain different aspects of 3D Earth structure. All kernels are computed with the help of adjoint techniques (e.g. Tarantola, 1988; Tromp *et al.*, 2005; Fichtner *et al.*, 2006a).

2.4 Sensitivity Kernel Gallery

To illustrate the main characteristics of our newly defined observables β_a and α_a , we present a gallery of sensitivity kernels for P, S, Love and Rayleigh waves. This is intended to aid in the development of the physical intuition necessary for the incorporation of β_a and α_a into tomographic inversions.

In the following figures we show sensitivity kernels for the apparent P and S velocities, $K_m(\alpha_a)$ and $K_m(\beta_a)$, but also for the velocity amplitude, $K_m(v)$, the rotation amplitude, $K_m(w)$, and the divergence amplitude, $K_m(s)$. This is because $K_m(\alpha_a)$ and $K_m(\beta_a)$ are equal to the differences

$$K_m(\alpha_a) = K_m(v) - K_m(s), \quad (2.6)$$

and

$$K_m(\beta_a) = K_m(v) - K_m(w). \quad (2.7)$$

Equations (2.6 and (2.7) follow directly from the product rule of differentiation. For a detailed derivation of equation (2.6) we refer to Fichtner & Igel (2009). Equation (2.7) follows analogously. The necessary technical details for the explicit computation of $K_m(v)$, $K_m(s)$ and $K_m(w)$ can be found in Appendix A.

2.4.1 P wave kernels

In our first series of examples we consider the direct P wave for the setup described in figures 2.1 and 2.2. Sensitivity kernels of the velocity amplitude v , the divergence amplitude s and the apparent P velocity α_a are shown in figure 2.4. All kernels are with respect to α , meaning that they describe

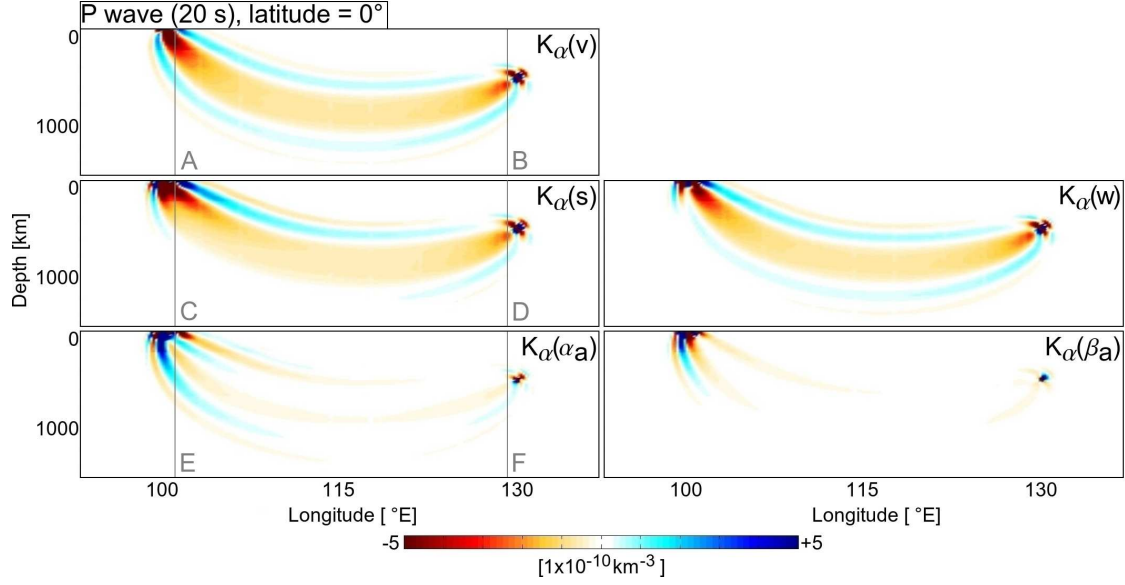


Figure 2.4: *Left column:* Vertical slices through the source-receiver plane of the sensitivity kernels for the velocity amplitude v , the divergence amplitude s and the apparent P velocity α_a . All kernels are relative to the P velocity α . $K_\alpha(\alpha_a)$ is most sensitive to near receiver structures. *Right column:* Vertical slices through the source-receiver plane of the rotation amplitude and the apparent S velocity kernels for the 20 s P wave from figure 2.2. $K_\alpha(\beta_a)$ is most sensitive to near receiver structures

the first-order response of the respective observable to a change in the P velocity. The velocity amplitude kernel $K_\alpha(v)$ has the typical cigar shape of a body wave amplitude kernel with negative sensitivity in the first Fresnel zone surrounding the geometric ray path (e.g. Dahlen & Baig, 2002). According to equation 2.5, positive perturbations of α located within the region of negative sensitivity reduce the P wave amplitude, and vice versa. The broad structure of the divergence amplitude kernel $K_\alpha(s)$ is similar to the velocity amplitude kernel $K_\alpha(v)$, meaning that both v and s provide nearly identical constraints on 3D Earth structure. As shown in figure 2.5, differences between $K_\alpha(v)$ and $K_\alpha(s)$ are mostly restricted to the near-receiver region and to the higher Fresnel zones. These differences are particularly evident in the apparent P velocity kernel $K_\alpha(\alpha_a)$, which, according to equation 2.6, is equal to the difference $K_\alpha(v) - K_\alpha(s)$. The localisation of $K_\alpha(\alpha_a)$ near the surface and the absence of a broad first Fresnel zone suggest that the apparent P velocity of the direct P wave constrains comparatively small-scale variations of α near the receiver. We furthermore note that $K_\alpha(\alpha_a)$ is predominantly positive, so that increases in α are expected to yield increases in α_a , and vice versa. The non-zero sensitivity of α_a directly at the source is confined to a very small volume, meaning that it is practically negligible.

While it is intuitively expected that the apparent P velocity α_a of the direct P wave is sensitive to the P velocity α , the behaviour of the apparent S velocity β_a is less predictable. First, we note that the apparent S velocity of the direct P wave takes a well-defined finite value. This is mostly

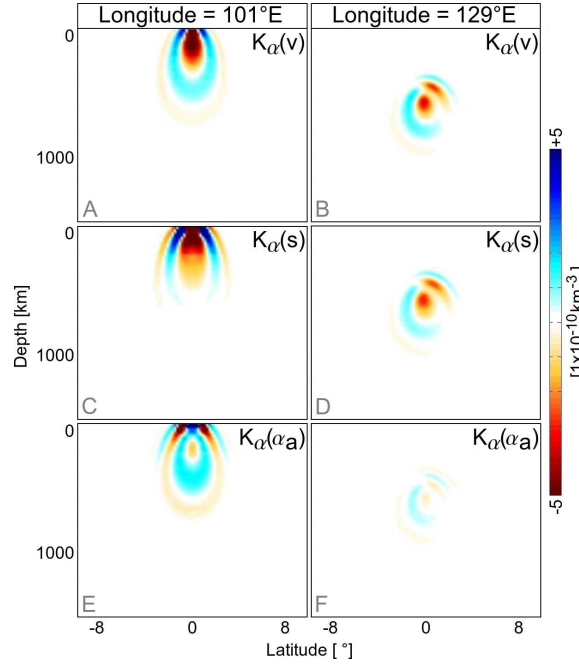


Figure 2.5: Vertical slices through various sensitivity kernels perpendicular to the source-receiver plane. The slices are close to the receiver in the left panel (A, C, E in figure 2.5) and close to the source in the right panel (B, D, F in figure 2.5). The velocity and divergence amplitude kernels, $K_\alpha(v)$ and $K_\alpha(s)$, are hardly distinguishable in the source region. Differences are most pronounced around the receiver. Consequently the α_a kernel $K_\alpha(\alpha_a) = K_\alpha(v) - K_\alpha(s)$ is largest in the receiver region (left) but vanishes closer to the source (right)

because the transverse rotation of the P wave (ω_θ in figure 2.2) is non-zero as a result of P-to-S conversions as the P wave reflects off the free surface. Changes in the P velocity α affect the P wave, and therefore also lead to perturbations of the converted S wave. This explains why the sensitivity of the rotational signal of the P wave, $K_\alpha(w)$, is significantly different from zero - and in fact very similar to the sensitivity of the displacement amplitude $K_\alpha(v)$ (figure 2.4). Again, the differences between $K_\alpha(v)$ and $K_\alpha(w)$ manifest themselves most clearly in the apparent S velocity kernel $K_\alpha(\beta_a)$, that is largest near the receiver, similar to $K_\alpha(\alpha_a)$.

It is a particularly noteworthy observation that the kernels $K_\alpha(\alpha_a)$ and $K_\alpha(\beta_a)$ from figure 2.4 are globally similar but differ significantly from each other in the near receiver region. This suggests that the apparent P and S velocities provide independent information on the near receiver P velocity structure. This leads us to conjecture that the combined use of both α_a and β_a in tomographic inversions can improve the resolution of 3D P velocity heterogeneity.

2.4.2 S wave kernels

The direct P waveform from the previous example is clearly separated from later-arriving phases, which allowed us to study unambiguously defined P wave kernels. S waveforms, in contrast, are

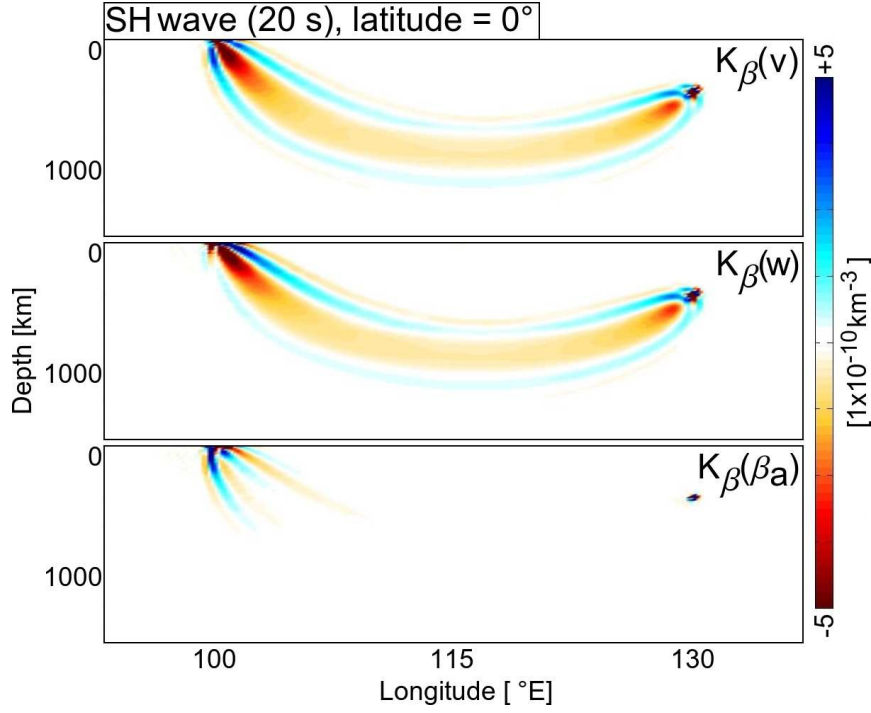


Figure 2.6: Vertical slices through the source-receiver plane of sensitivity kernels for the 20 s SH wave shown in figure 2.2. The velocity amplitude kernel at the top and the rotation amplitude kernel in the middle are similar. The apparent S velocity kernel $K_\beta(\beta_a)$ at the bottom pronounces the small differences between $K_\beta(v)$ and $K_\beta(w)$

more complex and appear in the form of an isolated peak only in the transverse velocity v_θ and the vertical rotation ω_z . Thus, in the interest of simplicity, we restrict our attention to the measurement of the apparent S velocity β_a , computed from the SH wave phase recorded in the v_θ and ω_z seismograms of figure 2.2.

The corresponding kernels for the rotation amplitude, $K_\beta(w)$, the velocity amplitude, $K_\beta(v)$, and the apparent S velocity, $K_\beta(\beta_a)$, are displayed in figure 2.6. All kernels are with respect to the S velocity β because the SH wave is practically insensitive to the P velocity α . Except for the slimmer first Fresnel zone, that results from the shorter wavelength of S waves compared to P waves, the S wave kernels duplicate the main features of the P wave kernels in figure 2.4. In particular, the large-scale features of $K_\beta(v)$ and $K_\beta(w)$ are nearly identical, meaning that the velocity amplitude and the rotation amplitude of the SH wave do not provide independent constraints on 3D S velocity heterogeneity.

As for the P wave, the essential benefit comes from the combination of the two measurements, v and w , into one new observable: the apparent S velocity β_a . The sensitivity of β_a to the S velocity, $K_\beta(\beta_a) = K_\beta(v) - K_\beta(w)$, highlights the differences between $K_\beta(v)$ and $K_\beta(w)$, that can mostly be found within the higher Fresnel zones and near the surface. The apparent S velocity kernel $K_\beta(\beta_a)$ is therefore - similar to the apparent P velocity kernels - only sensitive to compar-

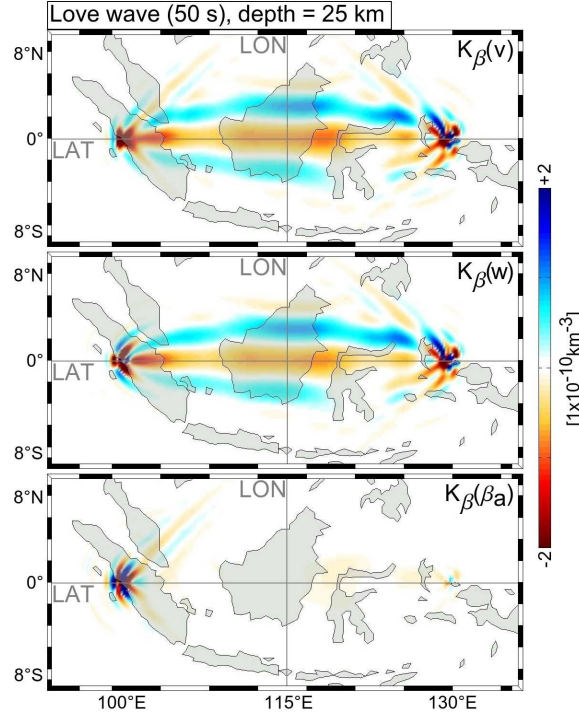


Figure 2.7: Horizontal slices at 25 km depth through the sensitivity kernels $K_\beta(v)$, $K_\beta(w)$ and $K_\beta(\beta_a)$ for a 50 s Love wave

atively small-scale near-receiver structure. It follows that β_a provides additional information on near-surface structure that is independent of the S velocity at larger distances from the receiver.

2.4.3 Love wave kernels

For the computation of surface wave sensitivity kernels we slightly modify our simulation setup: The source is placed at a shallower depth of 50 km, and the dominant period is increased to 50 s. In our study of Love waves we follow the previous SH wave example and consider only the v_θ and ω_z components. As Love waves are insensitive to the P velocity α , we concentrate on the sensitivity of β_a with respect to β .

Figure 2.7 demonstrates that the previously observed phenomenon of sensitivity restricted to the near-receiver region is reproduced by Love waves. It follows that Love waves should also be well suited to constrain small-scale variations in β close to the surface. Vertical slices through the sensitivity kernels $K_\beta(v)$, $K_\beta(w)$ and $K_\beta(\beta_a)$ are shown in figure 2.8. As expected, sensitivity for all measurements rapidly decreases away from the surface, and practically vanishes below 200 km depth.

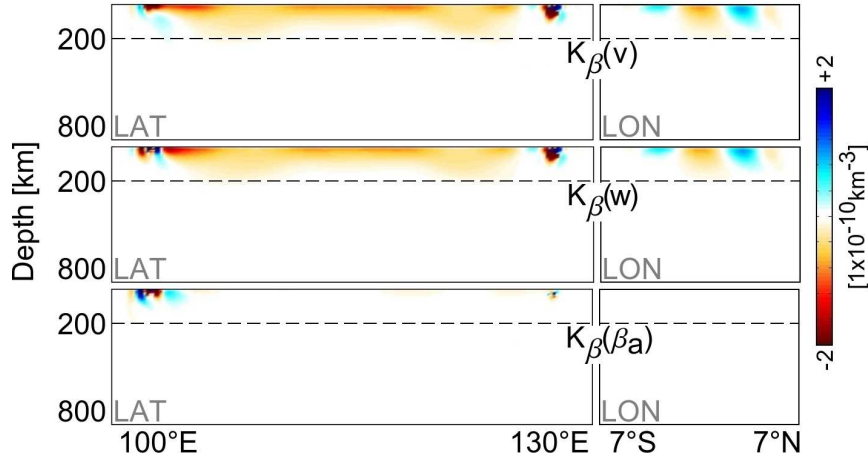


Figure 2.8: Vertical slices for a constant latitude at 0° (left panel) and a constant longitude at 115° (right panel) through the sensitivity kernels $K_\beta(v)$, $K_\beta(w)$ and $K_\beta(\beta_a)$ for a 50 s Love wave. The sensitivity of all measurements, v , w and β_a is restricted to the upper 200 km

2.4.4 Rayleigh wave kernels

To complete the gallery for the most prominent phases in a seismogram we consider 50 s Rayleigh waves. The setup is identical to the one used for Love waves. Fundamental Rayleigh waves are sensitive to the P velocity α primarily within the crust. We therefore focus on the sensitivity with respect to the S velocity β , displayed in figure 2.9. The velocity and divergence amplitude kernels are based on the v_ϕ and v_z velocity components, and the $e_{\phi\phi}$ and e_{zz} strain seismograms, respectively. The rotation amplitude kernel contains only the θ -component of the rotation seismogram.

Figure 2.9 reveals a phenomenon that is similar to the one encountered in section 2.4.1, where the rotation of the P wave was found to contain information on P velocity structure: The divergence of the Rayleigh wave is affected by heterogeneities in the S velocity, as evidenced by the large non-zero contributions to $K_\beta(s)$ in figure 2.9. As already expected the kernels for the apparent P and S velocities, $K_\beta(\alpha_a)$ and $K_\beta(\beta_a)$, are restricted to the vicinity of the receiver, with the most important contributions located in the higher Fresnel zones. Furthermore, the spatial patterns of $K_\beta(\alpha_a)$ and $K_\beta(\beta_a)$ differ strongly - as in the case of the P wave in section 2.4.1. This implies that β_a and α_a provide linearly independent constraints on 3D S velocity structure.

2.5 Perturbation Test

In the previous sections we studied the properties of various observables with the help of sensitivity kernels. The kernel $K_\beta(\beta_a)$, for instance, describes an *infinitesimal* change of the apparent S velocity, $\delta\beta_a$, in response to an *infinitesimal* S velocity perturbation $\delta\beta$. From a purely mathematical perspective, the kernel corresponds to an exact first derivative. Its physical meaningfulness, however, depends on the linearisability of the observable. How well does a kernel such as $K_\beta(\beta_a)$

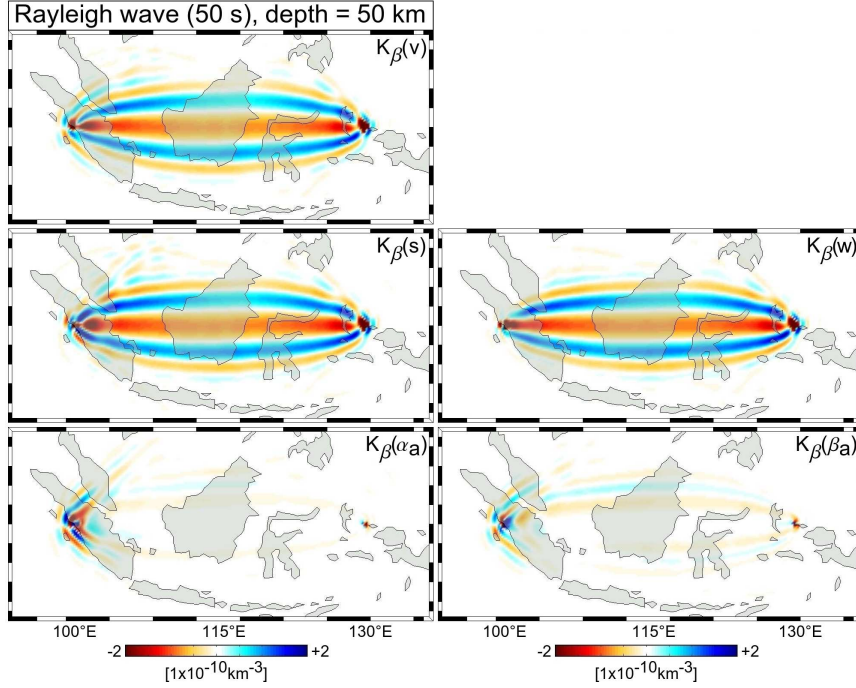


Figure 2.9: Horizontal slices at 50 km depth through the sensitivity kernels $K_\beta(v)$, $K_\beta(w)$, $K_\beta(s)$, $K_\beta(\beta_a)$ and $K_\beta(\alpha_a)$ for a 50 s Rayleigh wave

describe the *finite* change $\Delta\beta_a$ that results from a *finite* S velocity perturbation $\Delta\beta$?

To address this question we perform a simplistic perturbation test, summarised in figure 2.10 and figure 2.11. In terms of demonstrating precisely the nature of the new observables we abstain from more complex scenarios. For this we use the complete seismograms from the shallow event of sections 2.4.3 and 2.4.4 that are clearly dominated by surface waves. The synthetic waveforms are now recorded by a dense array of 720 equally distributed stations shown in figures 2.11a to 2.11c in the form of a regular mesh.

In the first simulation we compute synthetic velocity, rotation and strain seismograms for the 1D background model PREM (Dziewoński & Anderson, 1981). These provide reference values for the velocity amplitude, v^{ref} , the rotation amplitude, w^{ref} , and the apparent S velocity, β_a^{ref} . For the second simulation we add the $\pm 10\%$ chequerboard-like S velocity perturbation of figure 2.10 to PREM.

The resulting observables v^{pert} , w^{pert} and β_a^{pert} can then be used to compute the finite relative changes $(v^{\text{pert}} - v^{\text{ref}})/v^{\text{ref}}$, $(w^{\text{pert}} - w^{\text{ref}})/w^{\text{ref}}$ and $(\beta_a^{\text{pert}} - \beta_a^{\text{ref}})/\beta_a^{\text{ref}}$ at each station. These are displayed in figures 2.11a to 2.11c. The finite response of the velocity amplitude (figure 2.11a) corresponds well to the prediction of the sensitivity kernels $K_\beta(v)$ for both body and surface waves that are dominated by negative first Fresnel zones. An increase of β therefore leads to a decrease of v , and vice versa. This explains the approximate anti-correlation of $\Delta\beta$ and v in the vicinity of the perturbation. The large spatial extent of $K_\beta(v)$ is responsible for the significant changes

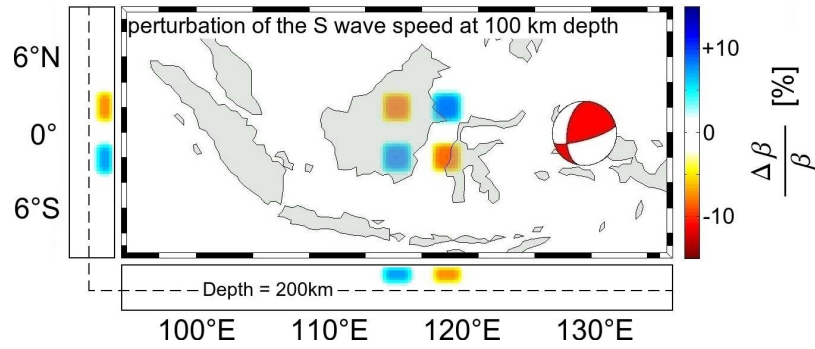


Figure 2.10: Checkerboard-like S velocity perturbation. The blocks are $2^\circ \times 2^\circ \times 200$ km wide, and are located directly beneath the surface. The perturbation amplitudes are max. ± 10 % relative to PREM

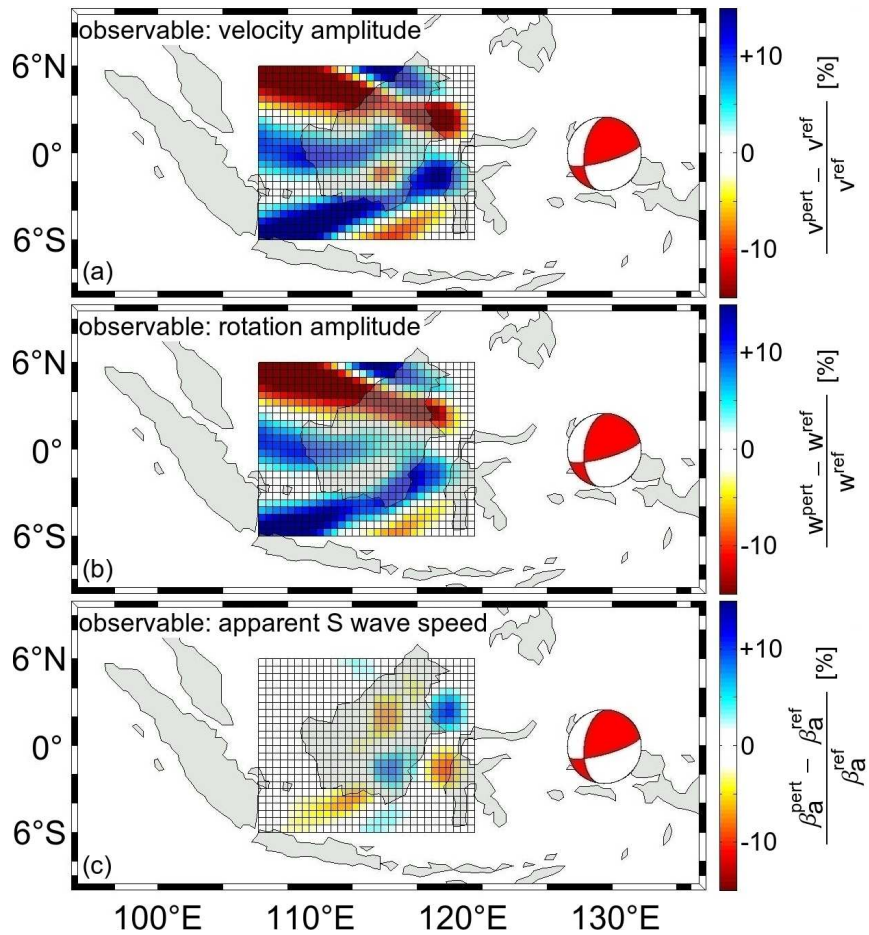


Figure 2.11: (a) Finite relative change of the velocity amplitude v across the array of 720 receivers. Each gridpoint represents one station. (b,c) The same as (a) but for the finite relative changes in the rotation amplitude w and the apparent S velocity β_a

in the velocity amplitude in regions that are far from the actual perturbation. In other words: the spatial pattern of $(v^{\text{pert}} - v^{\text{ref}})/v^{\text{ref}}$ does not allow us to unambiguously identify the location of the S velocity perturbation. The same arguments and conclusions are valid for the rotation amplitude w , shown in figure 2.11b.

The finite perturbations of the apparent S velocity, displayed in figure 2.11c, are localised directly above the S velocity perturbations, with only small tails extending towards the north- and south-west. This observation is in agreement with the concentration of the β_a kernels in the vicinity of the receiver. We may therefore directly infer both the location and the sign of lateral S velocity perturbations - which was, in fact, the initial motivation for defining the apparent S and P velocities.

While the previous example was designed to demonstrate the ability of the apparent S velocity to constrain local heterogeneities, a similar experiment is possible for the apparent P velocity, with nearly identical results.

2.6 Discussion

The sensitivity kernels for the apparent P and S velocities share two essential properties that are independent of the type of seismic wave considered: (1) the localisation of sensitivity in the vicinity of the receiver, and (2) the comparatively strong lateral variations of the kernels that result of the nearly complete absence of sensitivity inside the first Fresnel zone.

The advantageous consequence of the first property is that both α_a and β_a contain information about 3D heterogeneities in the direct vicinity of the receiver that is not contaminated by potentially unknown deeper Earth structure.

The rapid oscillations of the α_a and β_a kernels are both an advantage and a drawback. They suggest, on the one hand, that lateral variations smaller than the width of the first Fresnel zone may be resolvable. On the other hand, they are responsible for the small effect of larger-scale heterogeneities on the apparent P and S velocities. The relatively large-scale chequerboard-like pattern in figure 2.10, for instance, leads to nearly $\pm 15\%$ changes in v and w . However, the effect on β_a is almost a factor of 3 smaller, because the positive and negative contributions of $K_\beta(\beta_a)$ in the integral of equation 2.5 tend to cancel. Consequently high precision measurements of the new observables α_a and β_a are required. In this context providing reliable amplitude data is still a challenging task in seismic data acquisition.

The non-zero apparent S velocity of the P wave provides additional constraints on the local P velocity structure. This improvement - as well as its analogue for Rayleigh waves - should also be considered in the context of a multi-parameter inversion for both P and S velocity structure. Apparent velocities in the sense of equations (2.3) and (2.4) generally depend on both α and β . The solution of an inverse problem therefore requires either precise prior knowledge on one of the parameters (one-parameter inversion) or a joint inversion (multi-parameter inversion). In both cases, the benefits of incorporating apparent velocities in the inverse problem must be quantified with a

proper resolution analysis (Fichtner & Trampert, 2011)

The examples in this study are rather specific, but should be seen in a broader context. The apparent P and S velocities are just two among infinitely many combinations of translations, rotations and strain, that happen to have advantageous properties. One obvious extrapolation would be to construct combinations of measurements that are particularly useful to constrain specific aspects of 3D Earth structure. This could be achieved through the optimisation of a design criterion, e.g., large sensitivity or resolution in a certain region of the Earth. This, however, is clearly beyond the scope of this work.

2.7 Conclusions

The principal conclusions drawn from this work are as follows: (1) Both the apparent P and S velocities, α_a and β_a , are generally only sensitive to small-scale near-receiver structure - irrespective of the seismic wave considered. These properties result from the different source mechanisms of the adjoint wave fields for velocity, rotation and strain observables (see Appendix A). This suggests that measurements of α_a and β_a may be used in tomographic inversion to constrain local structure without requiring knowledge of 3D heterogeneities in the deeper Earth. (2) As a result of P-to-S conversions at the surface, the rotation of the P wave is significantly non-zero. This leads to a strong sensitivity of β_a to the local P velocity α . (3) The sensitivity patterns of $K_\alpha(\alpha_a)$ and $K_\alpha(\beta_a)$ for the P wave are substantially different, meaning that rotation measurements provide independent constraints on the local P velocity structure. (4) Similarly, α_a and β_a for the Rayleigh wave provide independent constraints on the local S velocity. (5) Perturbation tests confirm that finite perturbations in α_a and β_a are indeed well predicted by the sensitivity kernels that only capture the first-order effect. In particular, α_a and β_a only respond to near-receiver heterogeneities.

These results pave the way towards tomographic inversions for local Earth structure on the basis of combined translational, rotational and strain measurements.

Acknowledgments

We thank the members of the Munich Seismology group (LMU University, Munich) for many critical and fruitful discussions. The research presented in this article was supported by the International Graduate School THESIS within the Bavarian Elite Network. Andreas Fichtner was funded by The Netherlands Research Center for Integrated Solid Earth Sciences under project number ISES-MD.5. The numerical computations were performed on the National Supercomputer HLRB-II maintained by the Leibniz-Rechenzentrum. The constructive criticism of two anonymous reviewers allowed us to improve the first version of our manuscript.

Chapter 3

Optimal observables for multi-parameter seismic tomography

The content of this chapter is submitted as: Bernauer, M., Fichtner, A. and Igel, H., 12/2013. Optimal observables for multi-parameter seismic tomography. *Geophys. J. Int.*

Summary

We propose a method for the design of seismic observables with maximum sensitivity to a target model parameter class, and minimum sensitivity to all remaining parameter classes. The resulting optimal observables thereby minimise inter-parameter trade-offs in multi-parameter inverse problems.

Our method is based on the linear combination of fundamental observables that can be any scalar measurement extracted from seismic waveforms, including frequency-dependent traveltimes and amplitudes of arbitrary seismic phases. The optimal weights of the fundamental observables are determined with an efficient global search algorithm.

In a series of examples we illustrate the construction of optimal observables, and assess the potentials and limitations of the method. Combining measurements of Rayleigh wave traveltimes in only 4 frequency bands, yields an observable with strongly enhanced sensitivity to 3-D density structure. Simultaneously, sensitivity to S velocity (or shear modulus) is reduced, and sensitivity to P velocity (or bulk modulus) is practically eliminated. The original 3-parameter problem thereby collapses into a simpler 2-parameter problem with one dominant parameter.

By defining parameter classes to equal Earth model properties within specific regions, our approach mimics the classic Backus-Gilbert method where data are combined such that sensitivity is focused onto a part of the model that is of particular interest. This concept is illustrated using rotational ground motion measurements as fundamental observables. Forcing dominant sensitivity in the near-receiver region produces an observable that is completely insensitive to Earth structure

at more than a few wavelengths distance from the receiver. This observable may be used for local tomography with teleseismic data.

While our test examples use a small number of well-understood fundamental observables, few parameter classes and a radially symmetric Earth model, the method itself does not impose such restrictions. It can easily be applied to large numbers of fundamental observables and parameters classes, as well as to 3-D heterogeneous Earth models.

3.1 Introduction

3.1.1 Multi-parameter seismic inversion

The comprehensive characterisation of the Earth's internal state requires the combined knowledge of a broad spectrum of physical properties that may be inferred from seismic observations. Joint information on P velocity α , S velocity β and density ρ , for instance, is needed to independently constrain temperature and composition (e.g. Trampert *et al.*, 2004; Mosca *et al.*, 2012). Seismic anisotropy can be related to deformation (e.g. Hess, 1964; Gung *et al.*, 2003; Debayle & Ricard, 2013) and the presence of small-scale heterogeneity (e.g. Backus, 1962; Capdeville *et al.*, 2010; Fichtner *et al.*, 2012). The attenuation of seismic waves indicates high temperatures or the presence of fluids (e.g. Jackson *et al.*, 2002; Jackson, 2007; Dalton *et al.*, 2008), and lateral variations in density provide information on convective flow (e.g. Ishii & Tromp, 2001; Mosca *et al.*, 2012).

Inferring the state of the Earth from seismic observations is complicated because any observable d is sensitive to a wide range of properties, or parameter classes m_j . Arrival times of seismic waves, for example, are sensitive to both velocity and attenuation structure (e.g. Zhou, 2009; Ruan & Zhou, 2010). Wave amplitudes generally depend on attenuation and on purely elastic heterogeneities that cause focussing and defocussing (e.g. Dahlen & Baig, 2002; Zhou *et al.*, 2004). Shear wave splitting is affected by intrinsic anisotropy and small-scale heterogeneity, including fine lamination (e.g. Backus, 1962; Babuška & Cara, 1991).

All seismic inverse problems are *multi-parameter problems*, where trade-offs lead to uncertainties in and differences between seismic Earth models. The complete quantification of these trade-offs requires fully probabilistic inversions (e.g. Sambridge & Mosegaard, 2002; Bodin & Sambridge, 2009; Khan *et al.*, 2013) that are not feasible when the number of free parameters is large, or when the solution of the forward problem is computational intensive. While inter-parameter trade-offs are commonly recognised as a key problem, the design of deterministic (i.e. not probabilistic) multi-parameter inversions is mostly guided by subjective choices and technical convenience, instead of being driven by the physics of the problem. The most widely used approaches are (1) ignoring model parameters, (2) scaling, (3) sequential inversion, and (4) subspace methods.

Ignoring, for instance, the presence of 3-D heterogeneities in attenuation and P velocity is common in surface-wave tomography (e.g. Yoshizawa & Kennett, 2004; Fishwick *et al.*, 2005; Debayle &

Ricard, 2013). The neglected attenuation and P velocity structures may contaminate S velocity models.

To reduce the effective number of model parameters, scaling relations are also frequently applied. P wave anisotropy, for instance, may be scaled to S wave anisotropy in surface-wave tomography (e.g. Panning & Romanowicz, 2006; Nettles & Dziewoński, 2008) using, for instance, prior information from mineral physics (e.g. Montagner & Anderson, 1989). Similarly, density variations have been scaled to seismic velocity variations in joint inversions of gravity and seismic data (e.g. Tondi *et al.*, 2000, 2009; Maceira & Ammon, 2009). Scaling implements prior assumptions. It therefore precludes the detection of those interesting cases where the assumptions do not hold, e.g. when density and velocity heterogeneities are not correlated due to chemical heterogeneities.

Sequential inversion is widely used, for instance, in attenuation tomography. First, velocity variations are constrained using traveltimes information. Subsequently, the amplitudes of seismic waves are inverted for 3-D attenuation structure (e.g. Gung & Romanowicz, 2004; Kennett & Abdullah, 2011). While practical, this approach ignores that the velocity model will to some extent be incorrect because the traveltimes used for its construction have sensitivity to 3-D attenuation as well. Sequential inversion ignores that any observable depends on more than one parameter class, and it relies on the subjective decision that one parameter is more important than another, and therefore should be considered first.

Subspace methods for multi-parameter inversion (Kennett, 1997; Kennett *et al.*, 1988) introduce a scaling between different parameter classes that is determined by the data functional. While being preferable to a priori scalings, the use of subspace methods constitutes a subjective choice in itself, and it does not prevent the occurrence of inter-parameter trade-offs.

3.1.2 The origin of inter-parameter mappings in tomographic inversion

Inter-parameter mappings are an artefact of the regularisation needed to enforce a unique solution of otherwise ill-posed tomographic systems. This can be seen most easily for the example of a generic linear problem where variations in the observable data vector $\delta\mathbf{d}$ are related to variations in the model parameter vector $\delta\mathbf{m}$ via a sensitivity matrix \mathbf{G} ,

$$\delta\mathbf{d} = \mathbf{G} \delta\mathbf{m}. \quad (3.1)$$

In the unrealistic case where \mathbf{G} is invertible, $\delta\mathbf{m}$ is uniquely determined by $\delta\mathbf{d}$. The resolution matrix \mathbf{R} is then equal to the unit matrix, meaning that variations in one parameter class do not affect the inferred variations in another parameter class. When \mathbf{G} is not invertible, the solution of the tomographic system is often defined as the minimum of the quadratic misfit functional

$$\chi(\delta\mathbf{m}) = \frac{1}{2}(\delta\mathbf{d} - \mathbf{G} \delta\mathbf{m})^T \mathbf{C}_d^{-1} (\delta\mathbf{d} - \mathbf{G} \delta\mathbf{m}) + \frac{1}{2} \delta\mathbf{m}^T \mathbf{C}_m^{-1} \delta\mathbf{m}, \quad (3.2)$$

where \mathbf{C}_d and \mathbf{C}_m are the data and model covariance matrices, respectively. The normal equations that result from forcing the derivative of (3.2) with respect to $\delta \mathbf{m}$ to zero are

$$\delta \mathbf{m} = [\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_m^{-1}]^{-1} \mathbf{G}^T \mathbf{C}_d^{-1} \delta \mathbf{d} = \mathbf{L} \delta \mathbf{d}, \quad (3.3)$$

with the pseudo inverse $\mathbf{L} = [\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_m^{-1}]^{-1} \mathbf{G}^T \mathbf{C}_d^{-1}$. For the simplest case of only two parameter classes denoted \mathbf{m}_1 and \mathbf{m}_2 , we have

$$\mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \quad \mathbf{G} = (\mathbf{G}_1 \ \mathbf{G}_2), \quad (3.4)$$

and the pseudo inverse takes the specific form

$$\begin{aligned} \mathbf{L} &= [\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_m^{-1}]^{-1} \mathbf{G}^T \mathbf{C}_d^{-1} \\ &= \begin{pmatrix} \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_1 + \mathbf{C}_{m_1}^{-1} & \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_2 \\ \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_1 & \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2 + \mathbf{C}_{m_2}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{G}_1^T \\ \mathbf{G}_2^T \end{pmatrix} \mathbf{C}_d^{-1}. \end{aligned} \quad (3.5)$$

Ideally, the prior model covariances \mathbf{C}_{m_1} and \mathbf{C}_{m_2} reflect our - to some degree subjective - a priori guess on the plausible variability within the model parameter classes m_1 and m_2 , respectively. More often, however, \mathbf{C}_{m_1} and \mathbf{C}_{m_2} are chosen pragmatically, i.e., such that the matrix $[\mathbf{G}^T \mathbf{C}_d^{-1} \mathbf{G} + \mathbf{C}_m^{-1}]$ in equation (3.3) and (3.5) becomes invertible. This regularisation in the form of unavoidably non-zero model covariances - whether idealistic or pragmatic - introduces off-diagonal terms in the resolution matrix $\mathbf{R} = \mathbf{L} \mathbf{G}$, explicitly given by

$$\mathbf{R} = \begin{pmatrix} \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_1 + \mathbf{C}_{m_1}^{-1} & \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_2 \\ \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_1 & \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2 + \mathbf{C}_{m_2}^{-1} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_1 & \mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_2 \\ \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_1 & \mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2 \end{pmatrix}. \quad (3.6)$$

Off-diagonal elements in \mathbf{R} map variations in one parameter class into inferred variations of the other parameter class. Moreover, the regularisation imposed on $\delta \mathbf{m}_1$ affects $\delta \mathbf{m}_2$, and vice versa, thereby introducing additional interdependences that have no physical basis.

To avoid unphysical mappings, for instance, from \mathbf{m}_2 to \mathbf{m}_1 , the observables \mathbf{d} should ideally be chosen such that the elements of $\mathbf{C}_d^{-1/2} \mathbf{G}_2$ are much smaller than $\mathbf{C}_{m_2}^{-1}$, while keeping $\mathbf{C}_d^{-1/2} \mathbf{G}_1$ at order of $\mathbf{C}_{m_1}^{-1}$ or larger. Provided that such observables are available, the resolution matrix condenses to

$$\mathbf{R} = \mathbf{L} \mathbf{G} \approx \begin{pmatrix} [\mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_1 + \mathbf{C}_{m_1}^{-1}]^{-1} [\mathbf{G}_1^T \mathbf{C}_d^{-1} \mathbf{G}_1] & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad (3.7)$$

meaning that variations in \mathbf{m}_2 have little effect on the inferred variations in \mathbf{m}_1 . While illustrated here for the specific case of a linear least-squares problem, the phenomenon that necessary regular-

isation induces inter-parameter trade-offs, is general and methodology-independent. The challenge is therefore to find - or explicitly construct - observables that have maximum sensitivity with respect to \mathbf{m}_1 and minimum sensitivity with respect to \mathbf{m}_2 .

3.1.3 Outline

In the following paragraphs, we develop and apply a method for the design of seismic observables that reduce inter-parameter trade-offs and mappings by having maximum sensitivity with respect to the parameter class of interest while being as insensitive as possible to all other parameter classes. These *optimal observables* are defined as a linear combination of *fundamental observables* that are generally sensitive to multiple parameter classes. Possible choices of fundamental observables include, but are not limited to, finite-frequency traveltimes and amplitudes of body and surface waves, different types of waveform misfit in selected time intervals, or various measurements made on inter-station correlations of deterministic or quasi-random wave fields. The optimal weights in the linear combination are determined in a two-stage optimisation process that requires the sensitivity kernels of the fundamental observables.

This paper is organised as follows: In section 3.2, we develop a general method that allows us to design optimal observables on the basis of a collection of fundamental observables with less favourable properties. The theory will be followed by a series of examples where we design observables with maximum sensitivity to density (section 3.3.1), and maximum sensitivity in the near-receiver structure (3.3.2). These examples are primarily intended to illustrate the methodology, and to reveal its potentials and limitations. Finally, in section 3.4, we discuss various details of our method, including prerequisites for the successful construction of optimal observables, the role of prior model covariance and parameter scaling, as well as the use of optimal observables in iterative multi-parameter inversions.

In order to be independent of a specific inversion setup or discretisation, we will mostly work in the continuous space domain. The discrete versions of model parameters and sensitivity kernels can be obtained by projection onto an application-specific set of basis functions.

3.2 Theory

3.2.1 Constructing optimal observables

We start our development with a set of fundamental observables d_1, \dots, d_n , and a set of continuously distributed parameter classes $m_1(\mathbf{x}), \dots, m_p(\mathbf{x})$, where \mathbf{x} denotes the position vector. Fundamental observables can be, for instance, traveltimes in various frequency bands; and possible parameter classes include P velocity, S velocity and attenuation. Without loss of generality, we assume that the parameter class of interest corresponds to m_1 . Our goal is to find the linear combination of

fundamental observables

$$d = \sum_{i=1}^n w_i d_i, \quad (3.8)$$

that has maximum sensitivity to parameter class m_1 , while having minimum sensitivity to all other parameter classes m_2, \dots, m_p . To ensure well-posedness of this optimisation problem, we constrain the squared sum of the weights to equal 1, i.e.

$$\sum_{i=1}^n w_i^2 = 1. \quad (3.9)$$

Our task is to identify the appropriate vector of weighting coefficients $\mathbf{w} = (w_1, \dots, w_n)^T$. For this, we write the variation of the combined observable d ,

$$\delta d = \sum_{i=1}^n w_i \delta d_i, \quad (3.10)$$

in terms of the Fréchet or sensitivity kernels K_{ij} of the fundamental observables, defined as

$$\delta d_i = \sum_{j=1}^p \int_V K_{ij}(\mathbf{x}) \delta m_j(\mathbf{x}) d^3 \mathbf{x}, \quad (3.11)$$

with the Earth model volume denoted by V . Equation (3.11) is the space-continuous version of equation (3.1), with the sensitivity kernel K_{ij} playing the role of the sensitivity matrix \mathbf{G} . Inserting (3.11) into (3.10) gives

$$\delta d = \sum_{j=1}^p \int_V K_j(\mathbf{x}) \delta m_j(\mathbf{x}) d^3 \mathbf{x}, \quad (3.12)$$

where the sensitivity K_j of the combined observable with respect to parameter class m_j is simply the weighted sum of the fundamental sensitivities K_{ij} ,

$$K_j(\mathbf{x}) = \sum_{i=1}^n w_i K_{ij}(\mathbf{x}). \quad (3.13)$$

For d to be optimal, we require that the sensitivity power (Sieminski *et al.*, 2009)

$$P_j(\mathbf{w}) = \int_V K_j^2(\mathbf{x}) d^3 \mathbf{x} \quad (3.14)$$

is large for $j = 1$ and small for $j > 1$. We can cast this simultaneous maximisation and minimisation of sensitivity powers, into a single maximisation problem by defining the objective functional

$$\mathbb{P}(\mathbf{w}) = \sum_{j=1}^p b_j P_j(\mathbf{w}), \quad (3.15)$$

with balancing coefficients $b_1 > 0$ and $b_{j>1} < 0$. The balancing coefficients control the extent to which the sensitivity power P_1 is maximised while the sensitivity powers $P_{j>1}$ are minimised. Choosing, for instance, b_1 large relative to the $b_{j>1}$ will result in a combined observable d that has large sensitivity with respect to m_1 , but the sensitivities with respect to the remaining parameter classes $m_{j>1}$ may also still be large. On the other hand, choosing b_1 small relative to the $b_{j>1}$ will only minimise sensitivity with respect to the parameter classes $m_{j>1}$ without necessarily maximising sensitivity with respect to m_1 . Thus, appropriate balancing is needed, as we will explain in the following paragraphs.

Determining optimal weights w_i and balancing coefficients b_j simultaneously, is a non-linear problem that can be computationally expensive. To keep the computational burden at a manageable level, we developed a two-stage procedure for the solution of the optimisation problem:

Stage 1: We consider the balancing coefficients b_j constant, and we try to find weights w_i that maximise (3.15) subject to the constraint (3.9). The Lagrange function for this constrained optimisation problem is

$$\mathcal{L}(\mathbf{w}, \lambda) = \sum_{j=1}^p b_j P_j(\mathbf{w}) - \lambda \left(\sum_{i=1}^n w_i^2 - 1 \right). \quad (3.16)$$

Differentiating equation (3.16) with respect to \mathbf{w} , and using (3.13) and (3.14), yields

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}_l} = \sum_{j=1}^p b_j \int_V K_{lj}(\mathbf{x}) \left[\sum_{i=1}^n w_i K_{ij}(\mathbf{x}) \right] d^3 \mathbf{x} - \lambda \mathbf{w}_l, \quad (3.17)$$

with $l = 1, \dots, n$. Setting the left-hand side of (3.17) to zero, and defining the matrix

$$M_{li} := \sum_{j=1}^p b_j \int_V K_{lj}(\mathbf{x}) K_{ij}(\mathbf{x}) d^3 \mathbf{x} \quad (3.18)$$

with $l, i = 1, \dots, n$, condenses the optimisation problem into the eigenvalue equation

$$\mathbf{M}\mathbf{w} = \lambda \mathbf{w}. \quad (3.19)$$

All eigenvalues λ are real because \mathbf{M} is symmetric. Furthermore, we have $\mathbb{P}(\mathbf{w}) = \lambda$ for $\frac{\partial \mathcal{L}}{\partial \mathbf{w}_l} = 0$ (see Appendix B.1). Consequently, the solution of the maximisation problem is the weight eigen-

vector \mathbf{w} that corresponds to the maximum eigenvalue λ_{\max} of \mathbf{M} . The size of \mathbf{M} is controlled by the number of fundamental observables, which is typically small (in fact, around 5 in the examples shown later). It follows that the computational costs for solving the eigenvalue problem are generally low.

Step 2: So far, the solution vector \mathbf{w} of the maximisation problem still depends on the balancing vector $\mathbf{b} = (b_1, \dots, b_p)^T$ that we assumed fixed in equation (3.16), i.e. $\mathbf{w} = \mathbf{w}(\mathbf{b})$. As an additional optimisation step, the balancing coefficients may be further tuned to improve the properties of the combined observable d . For this, we draw test balancing vectors \mathbf{b}_{test} and evaluate the criterion

$$\frac{P_1[\mathbf{w}(\mathbf{b}_{\text{test}})]}{\prod_{j=2}^p P_j[\mathbf{w}(\mathbf{b}_{\text{test}})]} = \max, \quad (3.20)$$

which only involves the solution of the small eigenvalue problem (3.19). We then choose the balancing vector for which (3.20) attains its global maximum. In the applications of section 3.3, we generate test balancing vectors using a regular grid search.

We note that other criteria for finding optimal balancing coefficients may be used, depending on specific requirements and applications. Equation (3.20) proved useful for our purposes.

3.2.2 Parameter classes

The method developed in section 3.2.1 is very general, and does not pose any restrictions on what we define as a parameter class. In addition to physical properties such as P velocity and S velocity, parameter classes can also be defined as physical properties in specific regions of the Earth. In this sense, m_1 may, for instance, be the S velocity in the D'' region, and m_2 the S velocity in the remaining mantle and the crust. Using a suitable set of fundamental observables, our method may then be used to design observables with primary sensitivity to D'' structure. An example where sensitivity is focussed into a specific region, is shown in section 3.3.2.

3.3 Examples

To illustrate the construction of optimal observables, we provide various examples in a regional tomography setting. In section 3.3.1, we focus on traveltimes measurements used to construct combined observables with primary sensitivity to 3-D density structure. This will be followed, in section 3.3.2, by the combination of rotational and translational ground motion measurements into an observable that is sensitive only to near-receiver structure.

For our experiments we compute synthetic wave fields for a point source at 13.9 km depth beneath southern Greece. These are recorded at 20.98° epicentral distance in Malaga, Spain, as illustrated in figure 3.1 (left panel). In the interest of reproducibility, we use the 1-D Earth model ak135 (Ken-

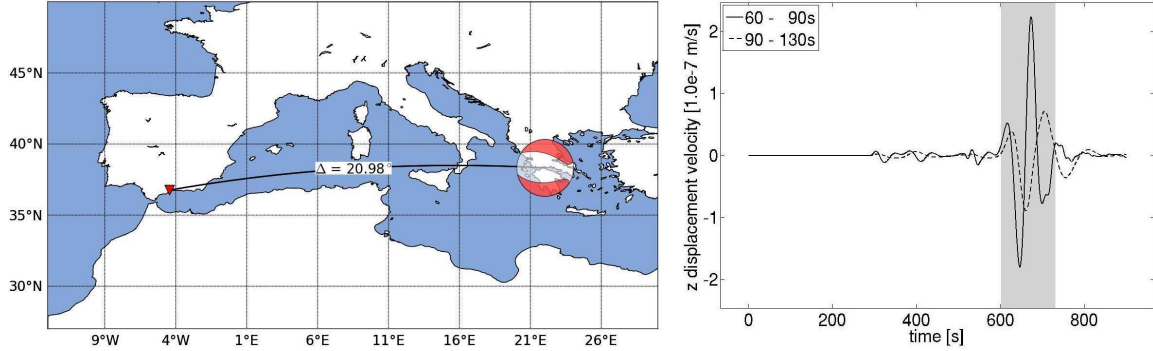


Figure 3.1: Left: source-receiver geometry for an event beneath southern Greece (longitude: 22.05° E, latitude: 38.32° N, depth: 13.9 km) and a receiver located in Malaga, Spain (longitude: 4.43° W, latitude: 36.76° N, depth: 0.0 km). The epicentral distance is 20.98° . The source mechanism is plotted at the epicentre. Right: vertical-component seismograms of the displacement velocity. The frequency bands are 60 – 90 s (solid line) and 90 – 130 s (dashed line). The Rayleigh wave window is marked by a grey bar.

nett *et al.*, 1995), and the source parameters correspond to the $M_w 5.5$ event in the southern Greece region on 2010/01/18 taken from the Global CMT catalogue (www.globalcmt.org). Velocity seismograms for the vertical component and frequency bands from 60 – 90 s (solid line) and 90 – 130 s (dashed line) are plotted in the right panel of figure 3.1.

3.3.1 Combining surface-wave traveltimes to enhance sensitivity to 3-D density variations

Trade-offs between parameter classes characterise inversions for 3-D density structure. While the Earth's mass, moment of inertia, and gravest normal modes constrain radial density structure to within $\sim 1\%$ over 200 km depth intervals (Kennett, 1998), 3-D density remains poorly known. Deterministic inversions are frequently biased by convenience-driven assumptions, including the neglect of compositional heterogeneities (e.g. Simmons *et al.*, 2010) and imposed correlations between velocities and density (Tondi *et al.*, 2000, 2009). Joint deterministic inversions of gravity and normal-mode data yield geodynamically plausible density heterogeneities for the lower mantle (Ishii & Tromp, 1999, 2001), but their reliability remains debated mostly because sensitivity to density is small and resolution estimates are biased by regularisation (e.g. Resovsky & Ritzwoller, 1999; Romanowicz, 2001; Kuo & Romanowicz, 2002). Probabilistic inversions for 3-D density avoid regularisation (e.g. Resovsky & Trampert, 2003; Mosca *et al.*, 2012) but are limited to smaller numbers of model parameters.

A central problem in density tomography is that the sensitivity of traveltime measurements to density variations tends to be small, oscillatory and parametrisation-dependent. In isotropic models parametrised in terms of P velocity α , S velocity β and density ρ , the traveltimes of body waves are nearly insensitive to density due to backward scattering (Wu & Aki, 1985; Tarantola, 1986).

While not being close to zero, the sensitivities of surface-wave traveltimes are vertically oscillating, meaning that positive and negative contributions of large-scale perturbations tend to cancel (e.g. Takeuchi & Saito, 1972). The alternative parametrisation in terms of bulk modulus κ , shear modulus μ and density, leads to large sensitivities of body- and surface-wave traveltimes to density, but also to strong inter-parameter trade-offs.

In the following paragraphs, we try to construct combinations of surface-wave traveltime measurements with dominant sensitivity to 3-D density variations. As fundamental observables we use cross-correlation time shifts of vertical-component surface waves in the period bands from 30-40 s, 40-60 s, 60-90 s, and 90-130 s, that we denote by $d_1 = d_{30-40}$, $d_2 = d_{40-60}$, $d_3 = d_{60-90}$, and $d_4 = d_{90-130}$. Thus, the number of fundamental observables is $n = 4$. The measurement time window is indicated by grey shading in the synthetic seismograms of figure 3.1. Sensitivity kernels needed in the optimisation scheme are computed with the help of adjoint techniques (e.g. Taran-tola, 1988; Tromp *et al.*, 2005; Fichtner *et al.*, 2006a; Liu & Tromp, 2008; Chen, 2011). Under the assumption that observed and synthetic waveforms are sufficiently similar, the sensitivity kernels for cross-correlation time shift measurements are independent of observed data (Luo & Schuster, 1991), meaning that we can construct optimal observables without requiring actual data.

Knowing that the inversion for 3-D density is a major challenge in seismic tomography, our goals are modest and mostly centred around a proof of the algorithmic concept introduced in section 3.2. Density is a difficult parameter that is well suited to illustrate the functioning of our method, as well as potential pitfalls and limitations. In a first numerical experiment, we construct an optimal observable with minimum sensitivity to α and β , and maximum sensitivity to the vertical density gradient $\partial_r \rho$ that can be related to the pressure derivative of density, observable in mineral physics experiments. This will be followed by a similar, though more difficult, attempt to find a optimal observable for density when the Earth model is parametrised in terms of the bulk modulus κ , the shear modulus μ and density. In section 3.3.1 we extend the analysis of a single source-receiver pair to a realistic configuration with multiple sources and receivers.

Optimal observable for the vertical density gradient $\partial_r \rho$

To construct optimal observables for the vertical density gradient $\partial_r \rho$, we start with the computation of sensitivity kernels for the relative perturbations $\delta \ln \alpha$, $\delta \ln \beta$ and $\delta \ln \rho$, where the reference values are taken from the 1-D Earth model ak135 (Kennett *et al.*, 1995). The kernels are displayed in the upper three rows of figure 3.2 for the frequency band from 60-90 s. The left panels show vertical slices through the source-receiver plane at 37.5°N . The right panel contains vertical slices at 9.0°E . While sensitivity with respect to $\delta \ln \alpha$ is non-zero only within the upper ~ 30 km, sensitivity to $\delta \ln \beta$ reaches its maximum around 80 km depth. The density kernel is characterised by a sign change around 50 km depth. It follows that density variations extending much more than 50 km in depth hardly affect 60 – 90 s surface-wave traveltimes because positive and negative contributions

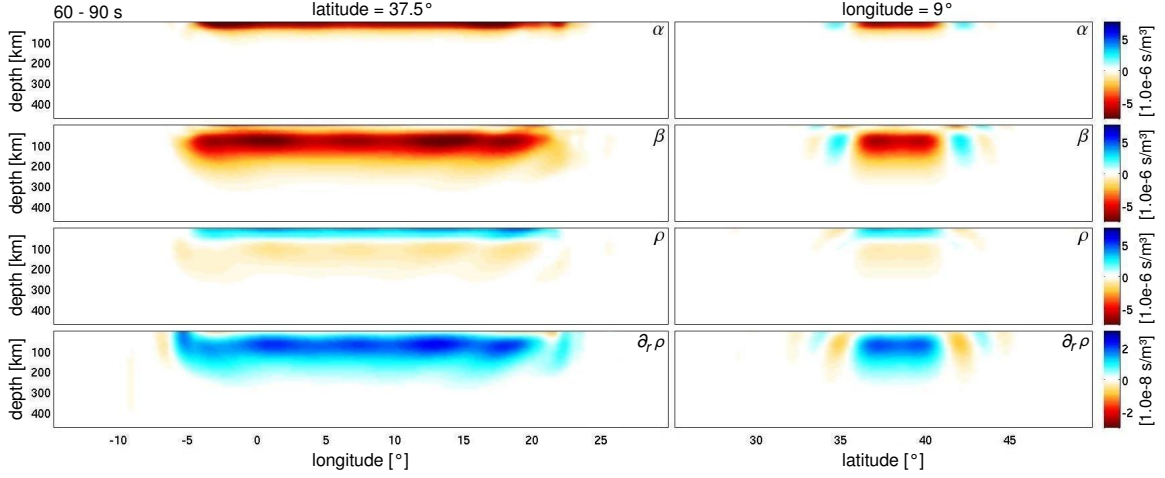


Figure 3.2: Vertical slices through the sensitivity kernels of a 60 – 90 s surface wave travel time with respect to the relative variations $\delta \ln \alpha$ (1st row), $\delta \ln \beta$ (2nd row) and $\delta \ln \rho$ (3rd row). The 4th row displays slices through the relative density gradient kernel, i.e. the kernel corresponding to $\delta \ln \partial_r \rho$.

tend to cancel. This characteristic is nearly period-independent.

As we demonstrate in Appendix B.2, the kernel for the vertical density gradient, $K_{\partial_r \rho}$, can be computed from K_ρ via a simple integral over radius from the center of the Earth ($r = 0$) to the surface ($r = R$):

$$K_{\partial_r \rho}(\theta, \phi, r) = - \int_{r'=0}^R K_\rho(\theta, \phi, r') dr'. \quad (3.21)$$

Kernels for relative perturbations in the vertical density gradient, $\delta \ln \partial_r \rho$, are obtained from $K_{\partial_r \rho}$ by multiplication with a reference value for $\partial_r \rho$. In the case of the 1-D reference Earth model ak135 (Kennett *et al.*, 1995), this reference takes the nearly constant value of $6.05 \cdot 10^{-4} \text{ kg/m}^4$ from 35-410 km depth. Vertical slices through the sensitivity kernel for $\delta \ln \partial_r \rho$ are shown in the fourth row of figure 3.2. Unlike K_ρ , the density gradient kernel $K_{\partial_r \rho}$ does not change sign with depth. It follows that variations of $\partial_r \rho$ extending over more than ~ 50 km in depth will affect the traveltimes observations.

Equipped with sensitivity kernels for $\delta \ln \alpha$, $\delta \ln \beta$ and $\delta \ln \partial_r \rho$ in the individual period bands, we can now use the machinery developed in section 3.2 to construct a composite observable that is optimally sensitive to the vertical density gradient. The resulting optimal weights are $w_1 = -0.11$, $w_2 = 0.41$, $w_3 = -0.73$, and $w_4 = 0.54$, meaning that, according to equation (3.8), the optimal observable takes the form

$$d = -0.11 d_{30-40} + 0.41 d_{40-60} - 0.73 d_{60-90} + 0.54 d_{90-130}. \quad (3.22)$$

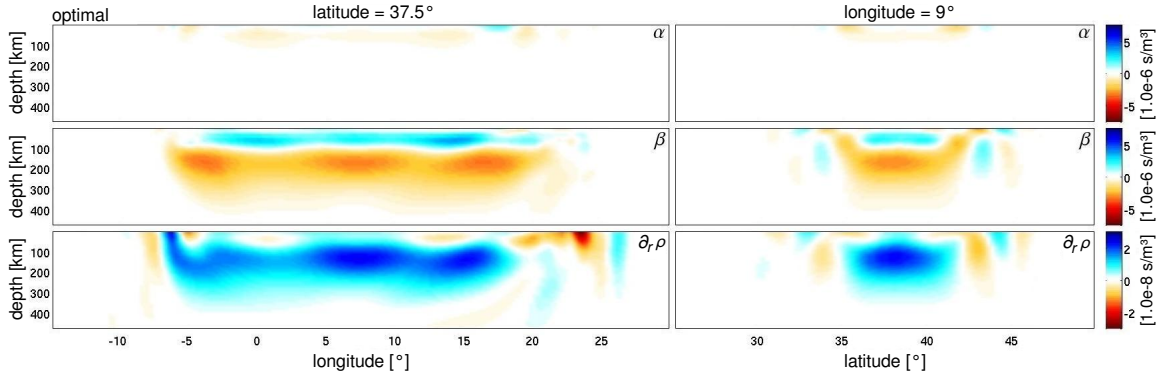


Figure 3.3: Vertical slices through the sensitivity kernels of the optimal observable d (equation (3.22)) with respect to the relative variations $\delta \ln \alpha$ (1st row), $\delta \ln \beta$ (2nd row) and $\delta \ln \partial_r \rho$ (3rd row). Sensitivity to P velocity is nearly eliminated, and sensitivity to S velocity is significantly reduced, relative to kernels for the individual period bands.

Sensitivity kernels corresponding to this optimal observable are displayed in figure 3.3. Compared to the 60 – 90 s period band kernels shown in figure 3.2, P velocity sensitivity is almost completely eliminated, and the maxima of S velocity sensitivity is reduced by $\sim 50\%$. Furthermore, the S velocity kernel of the optimal observable is oscillatory, meaning that large-scale variations in S velocity have an even smaller effect. At the same time, the density gradient kernel maintains the same amplitude as within the individual period bands.

A condensed version of the sensitivity kernels for the optimal observable is shown in figure 3.4 in the form of horizontally averaged kernels that emphasise their radial structure. While the sensitivity of the optimal observable relative to $\partial_r \rho$ consists of a single lobe with amplitude similar to the individual period band kernels, the sensitivities to α and β differ significantly from their individual contributions. Most importantly, and as seen in the vertical profiles of figure 3.3, sensitivity to α is eliminated almost completely, meaning that the original three-parameter inverse problem has reduced to a two-parameter problem that does not suffer from trade-offs with shallow P-velocity structure.

This first experiment indicates that the optimisation algorithm proposed in section 3.2 indeed produces observables that have increased sensitivity with respect to the target parameter, while having reduced sensitivity with respect to the remaining parameters. Remarkable aspects of this experiment are that relative sensitivity to density structure can be enhanced, and sensitivity to P velocity structure can be nearly eliminated by combining only a small number of Rayleigh wave travel-time measurements that by themselves have weak sensitivity to density and strong sensitivity to P velocity within the crust.

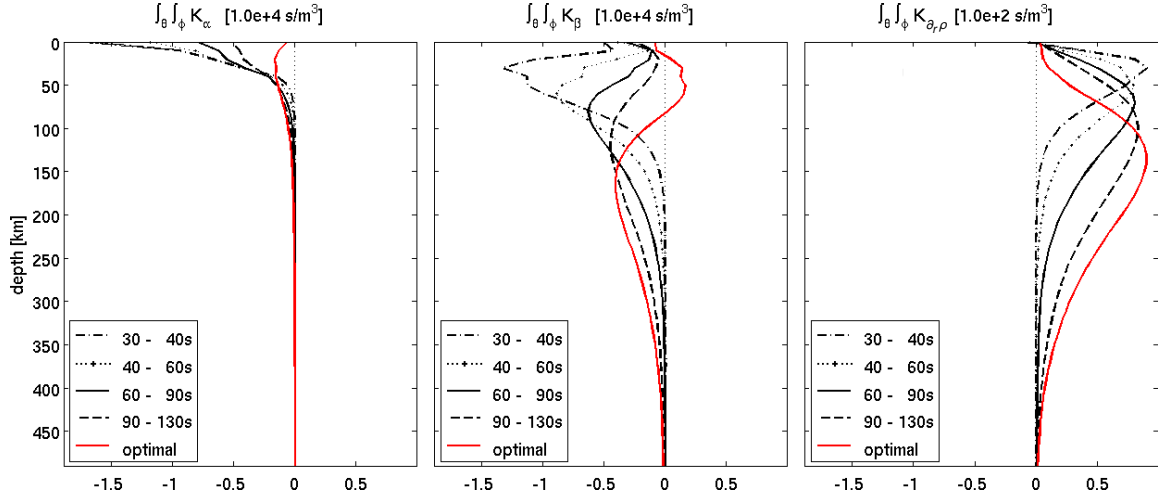


Figure 3.4: Horizontally averaged sensitivity kernels with respect to α (left), β (centre) and $\partial_r \rho$ (right). Kernels for the individual period bands are plotted in black. Red curves represent the kernels corresponding to the optimal observable. Sensitivity of the optimal observable to α is forced to almost zero, meaning that the original three-parameter problem has reduced to a two-parameter problem that is nearly independent of P velocity.

Changing parametrisation: $(\alpha, \beta, \rho) \rightarrow (\kappa, \mu, \rho)$

A complicating aspect of density tomography is the strong dependence of the density kernel on the choice of free parameters that can have profound effects on the result of deterministic regularised inversions (Cara *et al.*, 1984). The effect of changing the parametrisation from α , β and ρ to the bulk modulus κ , shear modulus μ and density, is visualised in figure 3.5 for the period band 60-90 s. While sensitivity with respect to κ is non-zero only in the upper ~ 30 km below the surface, sensitivity with respect to μ reaches its maximum at depths between 50 and 100 km and tends to zero towards the surface. The density kernel is characterised by an opposite sign relative to the κ and μ kernels, but there is no sign change with depth as in the α , β , ρ parametrisation.

Solving the optimisation problem produces the weights $w_1 = -0.28$, $w_2 = 0.6$, $w_3 = -0.65$ and $w_4 = 0.37$ that we use to construct the optimal composite observable. Vertical slices through and horizontal averages of the resulting sensitivity kernels are displayed in figures 3.6 and 3.7. Similar to the P velocity α in the previous example, sensitivity to κ is strongly reduced, meaning that the optimal observable is effectively sensitive to two parameters only. Sensitivity to μ , however, remains comparatively large. The failure to strongly reduce shear modulus sensitivity, results from the nearly period-independent geometry of the density kernels, the horizontal averages of which are shown in the right panel of figure 3.7. The small geometric variability reduces the ability to find linear combinations with fundamentally new properties, thereby illustrating a limitation of our method that we further discuss in section 3.4.1.

One of the main effects of the optimisation algorithm is to introduce additional sign changes with

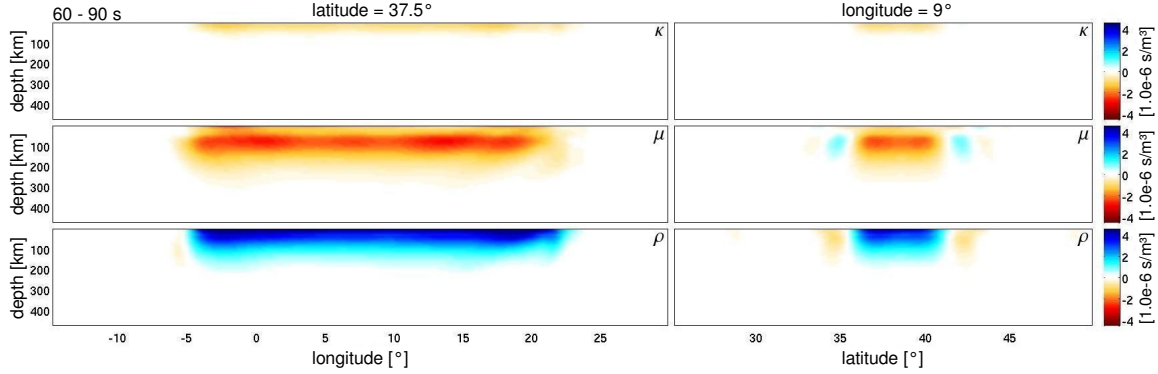


Figure 3.5: Cross-correlation time shift kernels with respect to the bulk modulus κ (first row), the shear modulus μ (second row) and density ρ (third row) for the frequency band 60 – 90 s. The vertical slices are shown at 37.5°N (left column) and 9.0°E (right column). The similarly shaped sensitivity kernels for μ and ρ indicate that strong trade-offs exist between both parameters.

depth in the sensitivity to μ . These sign changes reduce the impact of large-scale variations in μ on the optimal observable. In the following section we will further investigate this effect in a scenario with multiple sources and receivers.

A synthetic tomography scenario

In this section, we apply the construction of optimal observables to a configuration with multiple sources and receivers, as shown in figure 3.8. This experiment is intended to serve several purposes: (1) test the applicability of the optimisation algorithm to multiple source-receiver pairs, (2) check the reproducibility of the weighting coefficients w_i found for a single source-receiver pair in the previous example, and (3) test with actual measurements if density truly becomes the dominant parameter.

Our scenario includes ten shallow events in the Mediterranean region, with epicentres taken from the Global CMT catalogue (www.globalcmt.org). The positions of the 43 stations in the experiment correspond to station locations in the seismic networks IberArray (<http://iberarray.ictja.csic.es>) and ISIDe (<http://iside.rm.ingv.it/iside>). This distribution of sources and receivers ensures a dense coverage of the Mediterranean region except for the south-eastern part.

For the construction of optimal observables, we again consider cross-correlation time shifts of vertical-component Rayleigh waves for the frequency bands 30 – 40 s, 40 – 60 s and 90 – 130 s. We furthermore impose that the optimal weights w_i be the same for each receiver and event. For this scenario, the optimisation algorithm reproduces exactly the same weights as in the single source-receiver case of section 3.3.1. This result is plausible, at least for a radially symmetric background model where surface-wave sensitivity kernels have translational invariance and depth profiles that are nearly independent of epicentral distance. Differences in the source mechanisms

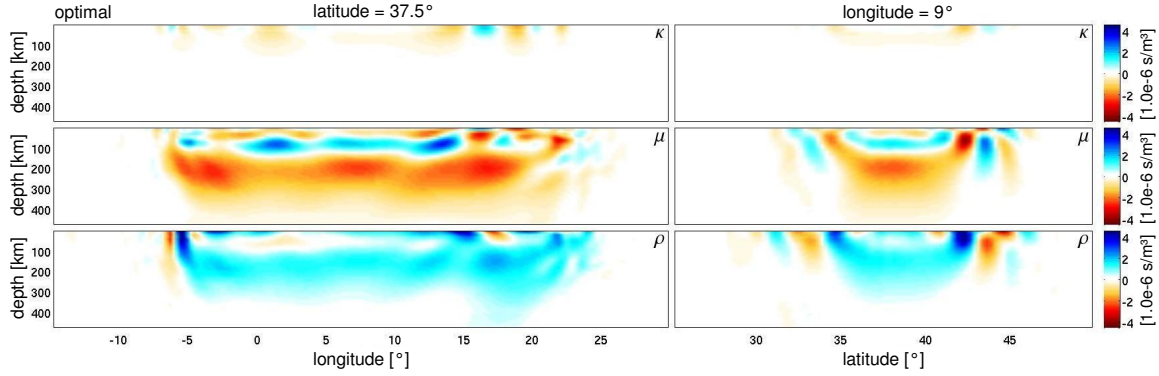


Figure 3.6: Vertical slices through sensitivity kernels for the optimal observable in the κ, μ, ρ parametrisation. While the sensitivity to κ is negligibly small, sensitivity to μ remains large. The comparatively small reduction of sensitivity to μ results from the small variability in the geomtry of the individual density kernels for different period bands, shown in the right panel of figure 3.7. The main effect of the optimisation is to introduce additional sign changes with depth in the sensitivity to μ .

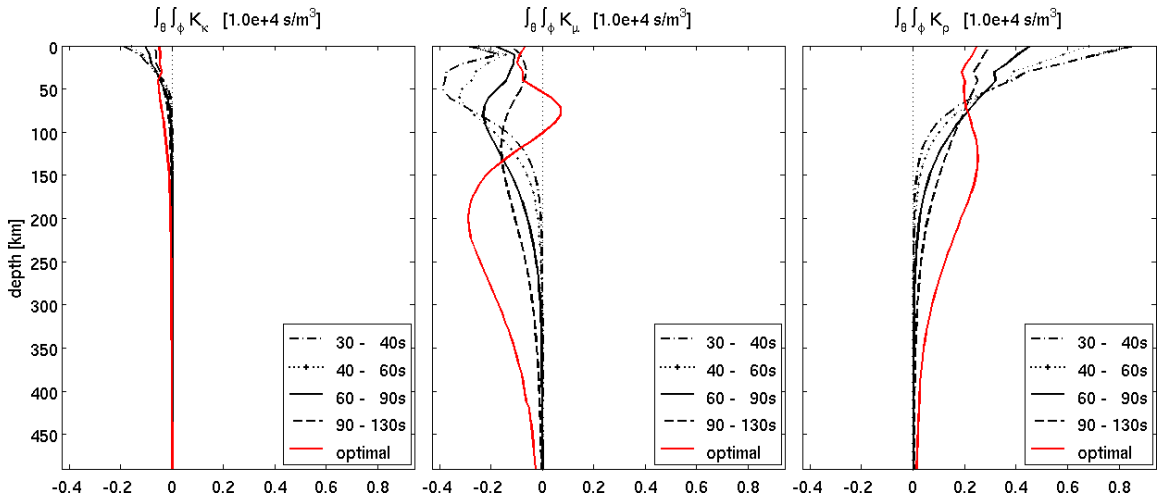


Figure 3.7: Horizontally averaged sensitivity kernels with respect to κ (left), μ (centre) and ρ (right). Kernels for the individual period bands are plotted in black. Red curves represent the kernels corresponding to the optimal observable. Sensitivity of the optimal observable to α is forced to almost zero. The comparatively small variability in the geomtry of the individual density kernels for different period bands limits the ability to find an optimal observable with a density sensitivity that is significantly increased relative to the shear modulus sensitivity.

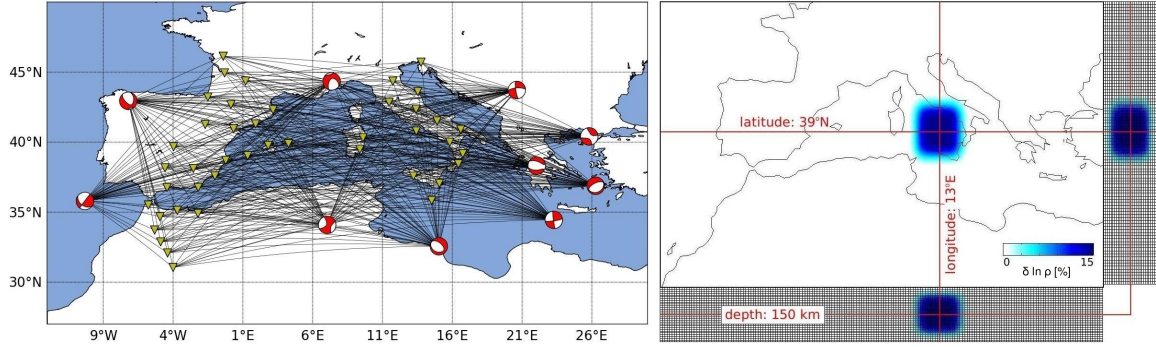


Figure 3.8: **Left:** Source/receiver geometry of the synthetic tomography scenario. Ten seismic events visualised by their corresponding source mechanisms encircle the Mediterranean Sea. The 43 receiver locations are marked by yellow triangles. Great circle paths are represented by black lines between each source/receiver pair. **Right:** Density perturbation of +15% centred in the Tyrrhenian Sea at 13°E, 39° N and 150 km depth.

play a minor role.

To test if the optimal observable indeed reacts primarily to density perturbations, we add heterogeneities to the 1-D model, centred at 13° E, 39° N and 150 km depth (see right panel of figure 3.8). In a first simulation we computed vertical-component seismograms for the individual frequency bands and a +15% density perturbation. For realistic Earth models, this value may be exaggerated, but it ensures that the traveltime differences between the 1-D model and its perturbed version are large and free of discretisation errors. In a second simulation we replaced the density perturbation by a −15% shear modulus perturbation. Following these simulations, we compute the cumulative rms time delays χ observed for all source-receiver pairs for both types of perturbations; and for the fundamental, as well as for the optimal observable.

The result is displayed in figure 3.9 in the form of relative time delays for density and shear modulus perturbations, $\chi(\delta \ln \rho) / \chi(\delta \ln \mu)$. Within the individual period bands, the time delays produced by the shear modulus heterogeneity is generally larger than the time delays produced by the density heterogeneity of comparable size. For the optimal observable, however, time delays due to the density heterogeneity are more than 40% larger, supporting the results of the sensitivity analysis of section 3.3.1 that already suggested the dominance of density structure. This result furthermore confirms that the optimisation algorithm produces the desired results, and indeed provides observables that mostly react to changes in the target parameter class.

3.3.2 Sensitivity optimisation involving translational and rotational ground motion measurements

So far, we were concerned with the optimisation of sensitivities with respect to a set of structural parameters. In the following application we demonstrate that our algorithm remains applicable when

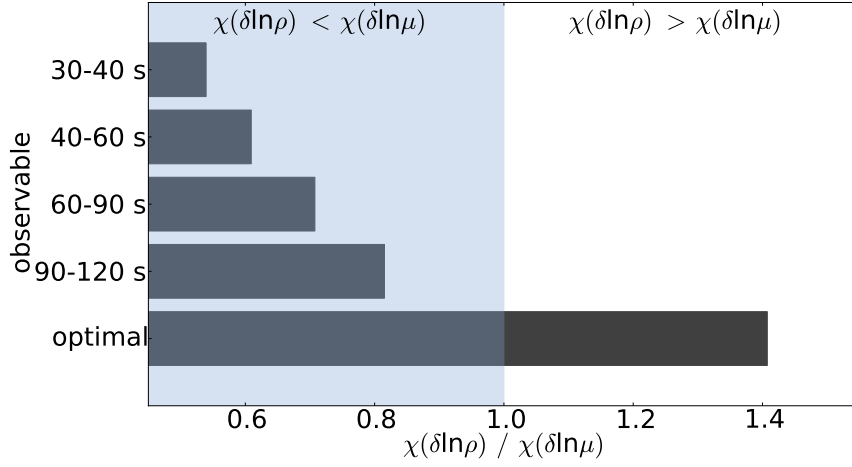


Figure 3.9: Comparison of cumulative time delays resulting from the density heterogeneity ($\chi(\delta \ln \rho)$) and the shear modulus heterogeneity ($\chi(\delta \ln \mu)$) shown in figure 3.8. The effect of the density heterogeneity is clearly dominant for the optimal observable.

the parameter classes are not structural parameters themselves, but structural parameters within certain target regions of the Earth.

For this demonstration, we consider the apparent S velocity β_a , defined at the receiver location \mathbf{x}^r as the ratio of rms translational velocity and rotation amplitudes of a seismic recording:

$$\beta_a(\mathbf{x}^r) := \frac{1}{2} \frac{\|\mathbf{v}(\mathbf{x}^r)\|_2}{\|\boldsymbol{\omega}(\mathbf{x}^r)\|_2}. \quad (3.23)$$

Quantities on the right-hand side of equation (3.23) denote the velocity amplitude $\|\mathbf{v}(\mathbf{x}^r)\|_2^2 = \int \mathbf{v}^2(\mathbf{x}^r, t) dt$ and the rotation amplitude $\|\boldsymbol{\omega}(\mathbf{x}^r)\|_2^2 = \int (\frac{1}{2} \nabla \times \mathbf{u})^2(\mathbf{x}^r, t) dt$. The quotient rule of differentiation implies that the sensitivity of β_a is equal to the difference between the sensitivities of $\|\mathbf{v}(\mathbf{x}^r)\|_2$ and $\|\boldsymbol{\omega}(\mathbf{x}^r)\|_2$:

$$K_\beta(\beta_a) = K_\beta(\|\mathbf{v}(\mathbf{x}^r)\|_2) - K_\beta(\|\boldsymbol{\omega}(\mathbf{x}^r)\|_2). \quad (3.24)$$

The key property of $K_\beta(\beta_a)$ is that sensitivity is concentrated in the immediate vicinity of the receiver, and tends to zero towards the source (Fichtner & Igel, 2009; Bernauer *et al.*, 2012). This property suggests a variant of local tomography with teleseismic data that is fully independent of both traveltime measurements and deep Earth structure far from the local target region (Bernauer *et al.*, 2009).

Our aim is to reproduce the localisation of sensitivity near the receiver by constructing an optimal observable, instead of using the definition of the apparent S velocity β_a from equation (3.23). For this, we define $d_v = \|\mathbf{v}\|_2$ and $d_\omega = \|\boldsymbol{\omega}\|_2$ as our fundamental observables. As first parameter class, we choose the S velocity β within a region R_1 that extends 500 km around the receiver.

The second parameter class is the S velocity in the remaining part of the model, denoted by R_2 . With the optimisation scheme developed in section 3.2, we construct an optimal observable $d = w_v d_v + w_\omega d_\omega$ with maximum sensitivity in R_1 and minimum sensitivity in R_2 .

In the practical implementation, we write the variation of d_v as

$$\delta d_v = \int_V G K_\beta(\|\mathbf{v}\|) \delta\beta d^3\mathbf{x} + \int_V (1 - G) K_\beta(\|\mathbf{v}\|) \delta\beta d^3\mathbf{x}, \quad (3.25)$$

where G is a Gaussian function that is centred at the receiver position and has a half-width of 500 km. Similarly, for the variation of d_ω , we have

$$\delta d_\omega = \int_V G K_\beta(\|\boldsymbol{\omega}\|) \delta\beta d^3\mathbf{x} + \int_V (1 - G) K_\beta(\|\boldsymbol{\omega}\|) \delta\beta d^3\mathbf{x}. \quad (3.26)$$

Equations (3.25) and (3.26) are special cases of equation (3.11) in the theoretical developments, and they imply that the four kernels entering the optimisation scheme are $G K_\beta(\|\mathbf{v}\|)$, $G K_\beta(\|\boldsymbol{\omega}\|)$, $(1 - G) K_\beta(\|\mathbf{v}\|)$ and $(1 - G) K_\beta(\|\boldsymbol{\omega}\|)$. These kernels are displayed in the first two rows of figure 3.10. All kernels were computed with the same setup as in section 3.3.1 but for a frequency band of 40-100 s.

Inserting the sensitivity kernels in the optimisation algorithm provides the weighting coefficients $w_v = 0.71$ and $w_\omega = -0.71$, meaning that the observable with maximum sensitivity in the receiver region and minimum sensitivity elsewhere is just the difference of d_v and d_ω , i.e.

$$d = 0.71 (d_v - d_\omega). \quad (3.27)$$

The factor $0.71 \approx \sqrt{1/2}$ results from the constraint that the squared sum of all weights be equal to 1 (see equation 3.9). The sensitivity kernel corresponding to d is given by

$$K_\beta(d) = 0.71 [K_\beta(\|\mathbf{v}(\mathbf{x}^r)\|_2) - K_\beta(\|\boldsymbol{\omega}(\mathbf{x}^r)\|_2)] = 0.71 K_\beta(\beta_a), \quad (3.28)$$

and is shown in the third column of figure 3.10. This result is remarkable for various reasons: First, the optimisation scheme indeed succeeds to focus sensitivity in the receiver region while completely erasing sensitivity in any other part of the Earth model. It follows that any shear velocity perturbation more than a few hundred kilometres from the receiver has no first-order effect on the optimal observable d . Second, we fully reproduce the sensitivity of the apparent S velocity β_a , up to a constant that results from the construction procedure of the optimal observable. Third, the result illustrates that kernels cannot be linked uniquely to an observable. Both $\beta_a = d_v/d_\omega$ and $d = d_v - d_\omega$ have, as a consequence of the quotient rule, identical sensitivity kernels.

Finally, we note that separating the complete kernels into low- and high-frequency components, instead of concentrating on different areas as in equations (3.25) and (3.26), also leads to exactly

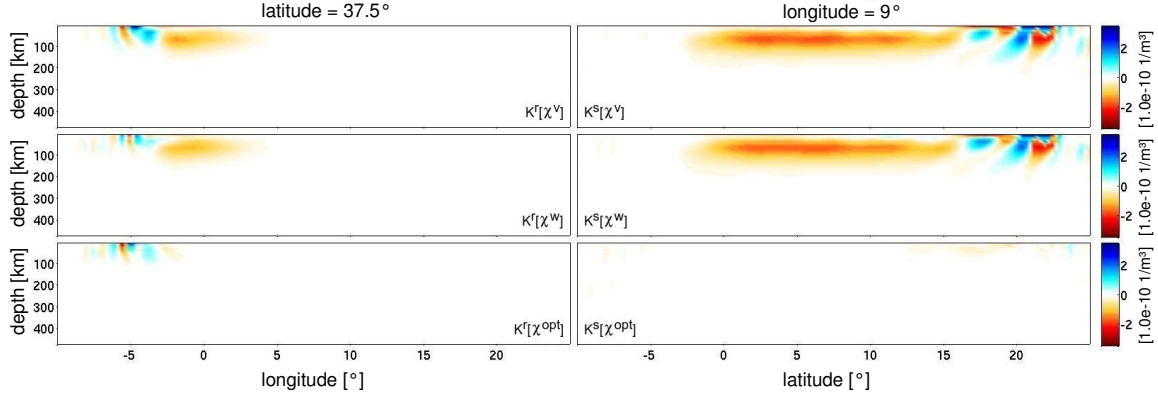


Figure 3.10: Vertical slices at 37.5°N through the sensitivity kernels $G K_\beta(\|\mathbf{v}\|)$ (top left), $G K_\beta(\|\boldsymbol{\omega}\|)$ (centre left), $(1 - G) K_\beta(\|\mathbf{v}\|)$ (top right) and $(1 - G) K_\beta(\|\boldsymbol{\omega}\|)$ (centre right). The function G is a Gaussian function that is centred at the receiver position and has a half-width of 500 km. The source-receiver geometry and the source characteristics are the same as in figure 3.1. The third row displays the kernels for the optimal observable $d = 0.71 (d_v - d_\omega)$. Shear wave sensitivity is entirely concentrated in the receiver region, and zero at distances of more than a few hundred kilometres from the receiver.

the same linear combination as in equation (3.27). The variations in d_v and d_ω are expressed as

$$\delta d_v = \int_V F * K_\beta(\|\mathbf{v}\|) \delta\beta d^3\mathbf{x} + \int_V (1 - F) * K_\beta(\|\mathbf{v}\|) \delta\beta d^3\mathbf{x}, \quad (3.29)$$

and

$$\delta d_\omega = \int_V F * K_\beta(\|\boldsymbol{\omega}\|) \delta\beta d^3\mathbf{x} + \int_V (1 - F) * K_\beta(\|\boldsymbol{\omega}\|) \delta\beta d^3\mathbf{x}, \quad (3.30)$$

where F is a spatial high-pass filter and $(1 - F)$ the corresponding low-pass filter. This result shows that the sensitivity close to the receiver in fact corresponds to the low-wavenumber contributions of the sensitivity kernels for d_v and d_ω .

3.4 Discussion

The series of examples shown in section 3.3 illustrates the applicability of the proposed method, but also raises various questions that deserve a more detailed discussion. These include the circumstances under which useful optimal observables can be designed, the role of quasi-subjective prior covariances, actual multi-parameter inversion schemes based on optimal observables, and the relation of our method to experimental design and the Backus-Gilbert method.

3.4.1 Successfully designing optimal observables

The extent to which sensitivity with respect to the target parameter class can be maximised while minimising sensitivity with respect to the remaining parameter classes depends critically on the shape of the sensitivity kernels for the fundamental observables. A large variability in the shape of the kernels generally favours the successful construction of an observable that is truly optimal in the sense of being sensitive to one parameter class only.

In the example of section 3.3.1, where we considered the sensitivity of Rayleigh wave traveltimes to the vertical density gradient $\partial_r \rho$, the shape of the individual kernels was strongly period-dependent. This dependence allowed us to reduce the sensitivity to α and β while keeping the sensitivity to $\partial_r \rho$ at a high level.

We encountered a less favourable situation when changing the parametrisation to κ , μ and ρ in section 3.3.1. In this parametrisation, the geometry of the density kernels is nearly period-independent, which limits the ability to find linear combinations with fundamentally different properties.

Along these lines, it is clear that the success of our methodology relies on the choice of fundamental observables. For the purpose of illustration, we deliberately chose a small number of fundamental observables, the physics of which are well understood. These observables allowed us to display both the functioning and limitations of our method. For future applications, the range of fundamental observables should clearly be extended, to include, for instance, time- and frequency-dependent amplitudes or various array-like measurements like inter-station correlations.

3.4.2 The role of prior model covariance and parameter scaling

The problem of inter-parameter trade-offs in regularised inversions has two components - one related to our prior assumptions on model covariance, and one related to the structural sensitivity of specific observables. The interplay of both components becomes most apparent in the generalised inverse, given in equation (3.5) for the case of two parameter classes.

Normalising, for instance, parameter class \mathbf{m}_1 by a very small prior standard variation σ_1 , will scale the corresponding \mathbf{G}_1 to $\sigma_1 \mathbf{G}_1$, meaning that the generalised inverse \mathbf{L} effectively becomes

$$\mathbf{L} \approx \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & [\mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2 + \mathbf{I}] \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{G}_1^T \\ \mathbf{G}_2^T \end{pmatrix} \mathbf{C}_d^{-1}. \quad (3.31)$$

The corresponding resolution matrix would be

$$\mathbf{R} \approx \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & [\mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2 + \mathbf{I}]^{-1} [\mathbf{G}_2^T \mathbf{C}_d^{-1} \mathbf{G}_2] \end{pmatrix}, \quad (3.32)$$

meaning that the prior assumption on model covariance causes $\delta \mathbf{m}_1$ to be completely unresolved. This illustrates the effect of the prior on perceived resolution in deterministic inversions; and there

is no generic solution other than probabilistic inversions that are, however, not feasible for high-dimensional model spaces.

Our development focusses on the second component of inter-parameter trade-offs that is related to the choice of observable. Independent of any prior assumptions, we ask how much sensitivity with respect to one parameter can be increased relative to sensitivity with respect to another parameter by designing optimal observables. The usefulness of these optimal observables in an actual regularised deterministic inversion also depends on the choices of prior covariances for the various parameter classes. Due to its subjective nature, this issue is not addressed here.

Empirically we found that a priori scalings of model parameters and sensitivities has little effect on the optimal observable. This is because the scaling of an individual sensitivity kernel K_j is compensated by the corresponding balancing factor b_j in equation (3.15).

3.4.3 Iterative inversion for multiple parameter classes

To use optimally designed observables in a tomographic inversion, we propose to proceed as follows: Again using the case of two parameter classes for illustration, we first construct an optimal observable \mathbf{d}_1 for \mathbf{m}_1 and a second optimal observable \mathbf{d}_2 for \mathbf{m}_2 . The linearised relation between observations and model parameters then takes the form

$$\begin{pmatrix} \delta \mathbf{d}_1 \\ \delta \mathbf{d}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \delta \mathbf{m}_1 \\ \delta \mathbf{m}_2 \end{pmatrix}. \quad (3.33)$$

In the ideal, but hardly achievable, scenario where \mathbf{d}_1 is insensitive to \mathbf{m}_2 and \mathbf{d}_2 is insensitive to \mathbf{m}_1 , the off-diagonal contributions \mathbf{G}_{12} and \mathbf{G}_{21} in equation (3.33) cancel, and the two parameter classes are decoupled. Under more realistic circumstances, \mathbf{G}_{12} and \mathbf{G}_{21} will be non-zero but by construction smaller than the diagonal elements \mathbf{G}_{11} and \mathbf{G}_{22} . To emphasise the dominance of the diagonal terms, we rewrite (3.33) as

$$\begin{aligned} \mathbf{G}_{11} \delta \mathbf{m}_1 &= \delta \mathbf{d}_1 - \mathbf{G}_{12} \delta \mathbf{m}_2, \\ \mathbf{G}_{22} \delta \mathbf{m}_2 &= \delta \mathbf{d}_2 - \mathbf{G}_{21} \delta \mathbf{m}_1. \end{aligned} \quad (3.34)$$

Following (Kennett & Sambridge, 1998), the first iteration consists in solving a regularised version of

$$\begin{aligned} \mathbf{G}_{11} \delta \mathbf{m}_1^{(1)} &= \delta \mathbf{d}_1, \\ \mathbf{G}_{22} \delta \mathbf{m}_2^{(1)} &= \delta \mathbf{d}_2. \end{aligned} \quad (3.35)$$

As a result of the enforced diagonal dominance, the first iterates $\delta \mathbf{m}_1^{(1)}$ and $\delta \mathbf{m}_2^{(1)}$ will already be useful approximations to $\delta \mathbf{m}_1$ and $\delta \mathbf{m}_2$. Subsequently, the off-diagonal terms are incorporated

iteratively by solving regularised versions of

$$\begin{aligned}\mathbf{G}_{11}\delta\mathbf{m}_1^{(i+1)} &= \delta\mathbf{d}_1 - \mathbf{G}_{12}\delta\mathbf{m}_2^{(i)}, \\ \mathbf{G}_{22}\delta\mathbf{m}_2^{(i+1)} &= \delta\mathbf{d}_2 - \mathbf{G}_{21}\delta\mathbf{m}_1^{(i)},\end{aligned}\tag{3.36}$$

with $i = 1, \dots$. During the inversion, only the right-hand side of the equations changes, meaning that it can be solved efficiently once generalised inverses of \mathbf{G}_{11} and \mathbf{G}_{22} are constructed.

3.4.4 Relation to the Backus-Gilbert method and experimental design

The linear combination of data for optimal-design purposes, proposed in equation (3.8), is reminiscent of the Backus-Gilbert approach for linear inverse problems with a single model parameter class (Backus & Gilbert, 1968). Backus and Gilbert suggested to combine data as $d = \sum_{i=1}^n w_i d_i$ such that the composite sensitivity kernel $K(\mathbf{x}) = \sum_{i=1}^n K_i(\mathbf{x})$ is close to $\delta(\mathbf{x} - \mathbf{x}_0)$, thereby producing a direct estimate of Earth model properties via $\delta d = \int K(\mathbf{x}) \delta m(\mathbf{x}) d^3\mathbf{x} \approx \delta m(\mathbf{x}_0)$. The concept is therefore similar, but the design goal and optimisation scheme differ significantly.

As indicated already in section 3.3.2, our method can be adapted to mimic the Backus-Gilbert method. Defining parameter class 1 to equal, for instance, S velocity at location $\mathbf{x} = \mathbf{x}_0$, and parameter class 2 to equal S velocity at all other positions $\mathbf{x} \neq \mathbf{x}_0$, would produce an optimal observable with an associated kernel that is as much δ -like as allowed by the data. While theoretically appealing, this approach would suffer from the same problems as the Backus-Gilbert method itself: To compute a whole-Earth model, the optimisation procedure needs to be repeated for all positions \mathbf{x} ; and the resulting Earth model may then not explain the data because the union of optimal point estimates is not necessarily an optimal collective estimate. This explains why applications of the Backus-Gilbert method are relatively few in number (e.g. Chou & Booker, 1979; Trampert & van Heijst, 2002).

Further links exist to experimental design frequently used to optimise the source-receiver geometry in geophysical exploration (e.g. Curtis & Maurer, 2000; Maurer *et al.*, 2010). Our method can be used for a similar purpose by defining the fundamental observables to be measurements for a collection of potential source-receiver pairs. In the final experiment, only those source-receiver pairs with a weight above a pre-defined threshold would then be used.

3.5 Conclusions

We developed a method for the construction of seismic observables that have maximum sensitivity with respect to a target model parameter class, while having minimum sensitivity with respect to all remaining parameter classes. This is achieved through the optimal linear combination of fundamental observables that can be any scalar measurement extracted from seismic recordings. The

resulting optimal observables minimise inter-parameter trade-offs that result from regularisation in ill-posed multi-parameter inverse problems.

In a series of examples we have shown that surface wave traveltime observations in different frequency bands can be combined such that sensitivity to 3-D density structure increases substantially. Simultaneously, sensitivity to S velocity (or shear modulus) is reduced, and sensitivity to P velocity (or bulk modulus) is practically eliminated, thereby reducing a 3-parameter problem into a simpler 2-parameter problem.

Using rotational ground motion measurements, allows us to construct an observable with dominant sensitivity in the near-receiver region, and zero sensitivity at more than a few wavelengths distance from the receiver. This observable is identical to the apparent shear velocity β_a (equation 3.23), originally defined on the basis of purely physical arguments (Fichtner & Igel, 2009; Bernauer *et al.*, 2012).

In the interest of reproducibility and clarity, our examples used a small number of well-understood fundamental observables, few parameter classes and a radially symmetric Earth model. The method itself, however, does not impose such restrictions, and it can easily be applied to large numbers of fundamental observables and parameters classes, as well as to 3-D heterogeneous Earth models.

Acknowledgments

The authors would like to thank Malcolm Sambridge and Jeannot Trampert for inspiring discussions and the provision of numerous useful references. We thank the members of the seismology groups in Munich (LMU University, Munich) and Utrecht (Universiteit Utrecht) for valuable discussions and pleasant working atmosphere. The research presented in this article was supported by the International Graduate School THESIS within the Bavarian Elite Network, the European Commission (Marie Curie Actions, ITN QUEST) and the Deutsche Forschungsgesellschaft (IG 16/15-1). The numerical computations were performed on the National Supercomputer SuperMUC maintained by the Leibniz-Rechenzentrum.

Chapter 4

Reducing non-uniqueness in finite source inversion using rotational ground motions

The content of this chapter is submitted as: Bernauer, M., Fichtner, A. and Igel, H., 02/2014. Reducing non-uniqueness in finite source inversion using rotational ground motions. *J. Geophys. Res.*.

Summary

We assess the potential of rotational ground motion recordings to reduce non-uniqueness in kinematic source inversions, with emphasis on the required measurement accuracy of currently developed rotation sensors. Our analysis is based on synthetic Bayesian finite source inversions that avoid linearizations and provide a comprehensive quantification of uncertainties and trade-offs. Using the fault and receiver geometry of the Tottori 2000 earthquake as a testbed, we perform inversions for two scenarios: In scenario I, we use translational velocity recordings only. In scenario II, we randomly replace half of the velocity recordings by rotation recordings, thus keeping the total amount of data constant. To quantify the noise-dependent impact of rotation recordings, we perform a sequence of inversions where the relative noise level of rotations and translations ranges between 0.5 and 5. Our results indicate that the incorporation of rotational ground motion recordings can significantly reduce non-uniqueness in finite source inversions, provided that measurement uncertainties are similar to or below the uncertainties of translational velocity recordings. When this condition is met, rupture velocity and rise time benefit most from rotation data. The trade-offs between both parameters are then strongly reduced, and the information gain nearly doubles. This suggests that rotational ground motion recordings may improve secondary inferences that rely on accurate information about rise time and rupture velocity. These include frictional properties of the

fault, radiation directivity and ground motion in general.

4.1 Introduction

4.1.1 Non-uniqueness in finite source inversions

The kinematics of earthquake sources, i.e. the evolution of the rupture process in space and time on a finite fault, are key to our understanding of seismic hazard and earthquake dynamics. The inverse problem of inferring a kinematic source model from seismic data is ill-posed, with solutions critically depending on (1) the parametrization of the model, (2) the inversion and regularization strategy, and (3) the available data (e.g. Olson & Apsel, 1982; Hartzell & Heaton, 1983; Custodio *et al.*, 2005; Hartzell *et al.*, 2007). Furthermore, non-linearities in the forward model, noisy data with insufficient coverage, and trade-offs between model parameters, result in a well-documented non-uniqueness of finite source models. Examples where different rupture scenarios explain observations equally well may be found, for instance, in Olson & Anderson (1988); Beresnev (2003); Ide *et al.* (2005), as well as in the Blindtest on Earthquake Source Inversion, initialized within the E.U. Project SPICE (Mai *et al.* (2007), <http://equake-rc.info/>). To reduce non-uniqueness in finite source inversions, various authors proposed to complement traditionally used strong-motion recordings by other data, most importantly including various types of geodetic measurements (e.g. Wald & Heaton, 1994; Hernandez *et al.*, 1999; Monelli *et al.*, 2009).

Over the past decade, reliable measurements of rotational ground motions have become available, in addition to the traditionally used translational ground motion. This new data type motivates the present study, which aims to assess the potential of incorporating rotational ground motion measurements to better constrain kinematic rupture models.

4.1.2 Rotational ground motions

Rotational ground motion naturally complements classically observed translational ground motion in E-W, N-S and vertical directions. Assuming infinitesimal deformation, the complete description of ground motion requires three components of translation, three components of rotation and six components of strain (e.g., Aki & Richards, 2002; Cochard *et al.*, 2006). Rotation ω is defined through a linear combination of space derivatives of the translation vector $\mathbf{u} = (u_x, u_y, u_z)$:

$$\begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \times \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \partial_y u_z - \partial_z u_y \\ \partial_z u_x - \partial_x u_z \\ \partial_x u_y - \partial_y u_x \end{pmatrix}. \quad (4.1)$$

Equation (4.1) suggests that the horizontal components of the rotation vector, known as tilt, effectively carry information on the seismic wave field at depth that is not available from conventional

arrays on the Earth's surface.

In the past decade, numerous studies established rotational ground motion measurements as a new source of information. Their potential field of application is multifaceted and involves geophysical exploration, hazard assessment and seismic engineering (e.g. Takeo, 1998; Wang *et al.*, 2009; Stupazzini *et al.*, 2009; Zembaty, 2009), as well as seismic tomography where the joint processing of translation and rotation data produces unique seismic observables with local sensitivities surrounding the receiver (Fichtner & Igel, 2009; Bernauer *et al.*, 2009, 2012).

However, the rotational part of ground motion has been ignored for a long time primarily because of the lack of suitable measurement devices. Recently developed high-resolution ring-laser gyroscopes (Schreiber *et al.*, 2009) produced the first consistent observation of teleseismic rotational ground motions in the vertical component, also known as torsion (Igel *et al.*, 2005, 2007). With the development of fiber-optic techniques, similar measurements for horizontal-component rotation are within reach (e.g. Bernauer *et al.*, 2012; Schreiber *et al.*, 2009; Velikoseltsev *et al.*, 2012).

4.1.3 Motivation and outline

Here we anticipate the availability of 6-C seismic arrays, i.e. arrays that provide three components of translation and three components of rotation data, and perform synthetic case studies to investigate the usefulness of such novel station networks for earthquake source inversion. The specific goals of this study are (1) to investigate if smaller 6-C arrays can provide more information on rupture processes than traditional 3-C arrays, and (2) to quantify upper bounds of the admissible noise level in rotation recordings relative to translation recordings, needed to obtain additional constraints on finite source parameters.

To avoid bias by subjective regularization, we base our analyses on a Bayesian, i.e. probabilistic, inversion approach (Sambridge & Mosegaard, 2002; Tarantola, 2005). This allows us to combine different data types in a natural way, and to quantify the information gain that results from the incorporation of rotational ground motions with different noise levels. Former probabilistic studies on earthquake source parameters (e.g. Peyrat & Olsen, 2004; Monelli *et al.*, 2009; Fichtner & Tkalčić, 2010) already demonstrated that probabilistic methods are well applicable to inherently ill-posed finite source inverse problems.

This paper is organized as follows: In section 4.2, we introduce the parametrization of the model that we use in this study including the event geometry, the distribution of stations and the representation of the fault plane. This leads to the formulation of the forward problem that defines the free parameters we invert for. As we primarily aim to study the impact of classic 3-C and novel 6-C data sets, we keep the rupture model and the forward solver deliberately simplistic. We describe the probabilistic source inversion approach in section 4.3, paying special attention to the assumed uncertainties in the data. Section 4.4 is dedicated to a series of synthetic case studies where we infer slip amplitudes, rupture velocity and rise time. For this, we use the source-receiver geometry

of the Tottori 2000 earthquake, for which various different rupture scenarios have been reported in the literature (e.g. Iwata & Sekiguchi, 2002; Semmane *et al.*, 2005; Piatanesi *et al.*, 2007; Monelli *et al.*, 2009). We analyze one scenario exclusively based on 3-C data and compare the results to a second scenario including 6-C data at only half as many stations. We conclude with an assessment of the accuracy requirements of rotational ground motion sensors used in finite source inversion.

4.2 Kinematic fault rupture modeling

In our synthetic source inversion experiments, we explore if 6-C seismic networks provide more information and thus reduce the non-uniqueness in finite source inversion compared to scenarios with only conventional 3-C arrays. To model a realistic test case, we implement the geometry of the M_w 6.7 Tottori, Japan, earthquake that occurred on 6 October 2000.

4.2.1 Scenario and event geometry

As the first major earthquake recorded by the Japanese K-net and KiK-net strong-motion network (<http://www.kik.bosai.go.jp/>), the Tottori 2000 earthquake is well documented and intensively studied. The event exhibits a comparatively simple fault geometry, inferred mainly from aftershocks (Fukuyama *et al.*, 2003). The focal mechanism is described by a left-lateral strike slip with a vertically dipping fault plane (dip = 90° , rake = 0°) striking at 150° , as illustrated in Figure 4.1a. Following the studies of Semmane *et al.* (2005) and Monelli & Mai (2008), we locate the hypocenter at 35.269°N , 133.357°E and 12.5 km depth; and we assume a buried fault plane with its upper edge at 2.75 km depth, 12 km dip length and 32 km strike length. Fixing these parameters a priori, we do not invert for hypocenter coordinates, the moment tensor, and the geometry of the fault plane. The synthetic data used in the test is simulated at 20 sites corresponding to the locations of K-/KiK net stations. The maximum receiver distance from the epicenter is 60 km (Figure 4.1a).

4.2.2 The forward problem

To kinematically describe the fault rupture evolution in time and space, we adopt the forward modeling approach used, for instance, by Cotton & Campillo (1995), Hernandez *et al.* (1999) or Semmane *et al.* (2005). For this, we divide the fault plane into 24 quadratic subfaults with a side length of 4 km. Each subfault consists of 64 regularly spaced moment tensor point sources, thus resulting in a grid spacing of 0.5 km (Figure 4.1b). We compute the 6-C response (velocities and rotation rates) up to a frequency of 1 Hz for each point source. For the wavefield simulations we use the 1-D purely elastic Earth model proposed by Semmane *et al.* (2005) and also used by the Research Center for Earthquake Prediction (RCEP) of the Disaster Prevention Research Institute (DPRI), Kyoto (see Appendix C.1). To be flexible for potentially more complex models in future studies, we applied a spectral-element solver to compute the seismic wavefield for the moment tensor point sources

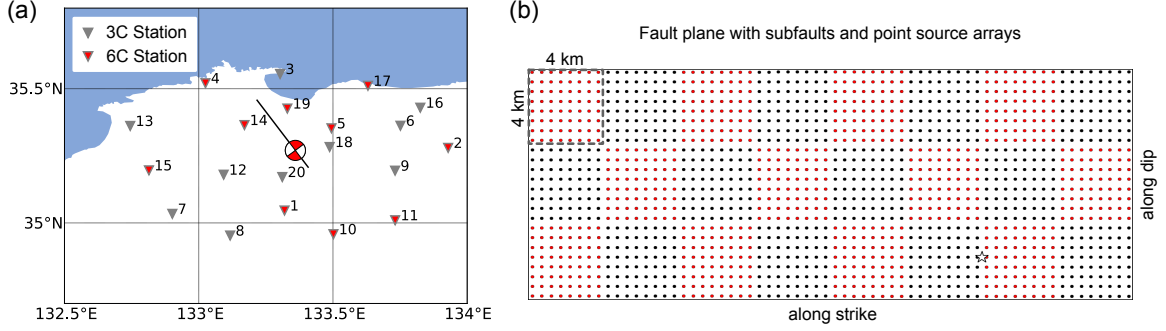


Figure 4.1: (a) Event geometry of the synthetic source inversion experiment based on the Tottori 2000 earthquake. Translational ground motions are simulated at all 20 stations (gray triangles). Triangles with a red marker symbolize a selection of 10 stations where both, synthetic translation and rotation data is considered. The fault strike is described by the black solid line. The focal mechanism plotted at the epicenter location corresponds to a left-lateral strike slip event. (b) Vertically dipping fault plane divided into 24 quadratic subfaults with a side length of 4 km. Each subfault is represented by a regular pattern of 64 moment tensor point sources. The white star locates the hypocenter.

on each subfault (Fichtner *et al.*, 2009). With a minimum wavelength of 3,179 m this set-up is similar to the previous kinematic source studies by Cotton & Campillo (1994, 1995) who propose a subfault length at the order of the shortest wavelength, with at least 6 point sources in vertical and horizontal direction. Fixing a local horizontal rupture velocity of 2500 m/s on each subfault, we compute a library with the seismic response of each subfault by summing over the previously computed point-source seismograms in each subfault, appropriately delayed in time. This library has to be computed only once and contains the seismic wavefield, G_{kl}^r , corresponding to subfault k , station r and component l .

We allow each subfault to slip once and parameterize the slip process in terms of slip amplitude ($slip_k$), rupture velocity (c^{rup}) and rise time (R). The slip amplitude is heterogeneous across the fault plane, leading to 24 free parameters. Together with the distance between the center of subfault k and the hypocenter, the rupture velocity provides the rupture time $t_k(c^{rup})$ of subfault k . The rise time expresses the duration of the slip. Both rupture velocity and rise time are homogeneous parameters across the fault plane. Thus, we invert for 26 free parameters in total. Finally, the complete seismic response, $v_l^r(\omega)$, at station r , component l and for the circular frequency, $\omega = 2\pi f$, is computed as a linear sum of $N(= 24)$ subfault contributions

$$v_l^r(\omega) = \sum_{k=1}^N slip_k \exp[-i\omega t_k(c^{rup})] G_{kl}^r(\omega) S(R, \omega). \quad (4.2)$$

In equation (4.2) S represents the source function that we implemented as an ordinary ramp function. Additional details on the source function are provided in Appendix C.2.

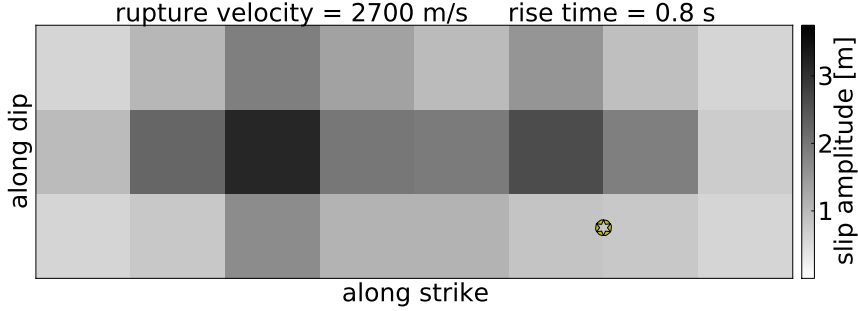


Figure 4.2: Heterogeneous slip amplitude distribution of the target model. Two slip amplitude peaks of 3.2 m and 2.6 m are centered in the left and the right half of the rupture plane. Rupture velocity (2700 m/s) and rise time (0.8 s) are homogeneous across the fault plane. The hypocenter is marked by a white star.

4.2.3 Target model and synthetic data

Performing a synthetic study means that we know the true kinematic rupture model (target model) we invert for in advance. Mimicking the results of former studies on the Tottori 2000 event (e.g. Semmane *et al.*, 2005; Monelli & Mai, 2008), we assume a heterogeneous distribution of slip amplitudes with two high-slip regions (asperities), as visualized in Figure 4.2. As realistic values for rupture velocity and rise time we use 2700 m/s and 0.8 s, respectively. Inserting the target model into equation (4.2), yields the synthetic data that we try to fit in the inversion process. Figure 4.3 shows some characteristic features of the synthetic data produced by the target model. For each station, the energy ratios, i.e. the L_2 -norm ratios of the seismograms, observed in fault parallel (gray), fault normal (black) and fault vertical (orange) components are displayed. Figure 4.3a corresponds to velocity and Figure 4.3b to rotation rate seismograms, respectively. Different values of the absolute energy measured at individual stations are encoded in the circle diameter at each station and are explained mainly by the station distance to the fault and the focal mechanism. As expected for a strike-slip event, Figure 4.3 clearly highlights that velocity measurements provide the highest energy in the horizontal components. For example at station 6, 7, 10, 15 or 16 the energy measured in the vertical components almost tends to zero. For rotation rates, the highest energy is mostly found in the vertical components, but there are also stations with considerable energy in tilt as for example station 1, 9, 14, 19 and 20. This concentration of energy in the horizontal velocity and vertical rotation rate seismograms is distinctive for a strike-slip event.

To render our synthetic study more realistic, we add Gaussian noise to each synthetic seismogram. Since we limit our analysis to a comparison of finite source inversions with and without rotational ground motions, only the relative noise level in translational and rotational seismograms is relevant. In our initial examples, we set equal noise levels in translations and rotations, with an amplitude of 1.0 % of the maximum amplitude observed in all velocity and rotation rate synthetics, respectively. This choice ensures that a potential gain (or loss) of information from any type of data does not

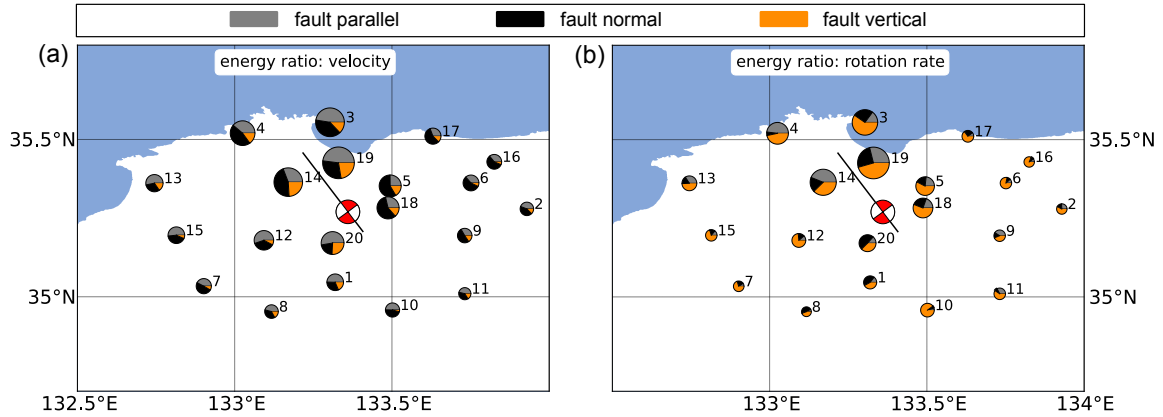


Figure 4.3: Energy ratios corresponding to fault parallel (gray), fault normal (black) and fault vertical (orange) component seismograms. The circle diameter reflects the absolute energy values in the synthetic data for each station. (a) Concerning velocity seismograms the highest energy is observed in the horizontal components. (b) Rotation rates provide the highest energy mostly in the vertical components even though at several stations considerable energy is present also in the horizontal components.

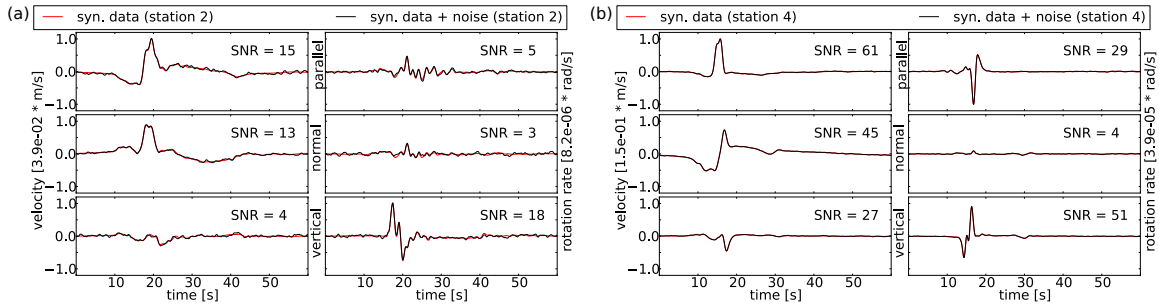


Figure 4.4: Synthetic data (red) and synthetic data perturbed with Gaussian noise (black). (a) Fault parallel, normal and vertical seismograms and their corresponding signal-to-noise ratio (SNR) are plotted for velocity (left panel) and rotation rate (right panel) observations at station 2. (b) The same as (a) but for station 4.

result from different noise levels. In section 4.4.4 we perform studies with varying noise levels in the rotation recordings, thus exploring the requirements on measurement accuracy.

Figure 4.4 demonstrates how the additional Gaussian noise affects the synthetic data. Station 2 (Figure 4.4a) is one of the large-distance stations and represents low amplitude data. The signal-to-noise ratio (SNR), i.e. the ratio between maximum data and noise amplitudes, varies between 3 and 18, depending on the component. For stations closer to the fault, the noise is negligible, as illustrated for station 4 in Figure 4.4b.

4.3 Probabilistic finite source inversion

Finite source inversion is ill-posed due to the sparsity of data, the presence of noise, and the non-linear relationship between model parameters and data.

Deterministic source inversion attempts to find the model that best explains the data, typically using iterative optimization techniques. The lack of information with its resulting non-uniqueness is compensated by regularization, e.g. in the form of smoothing and damping. Since regularization is to some degree subjective, resolution estimates tend to be biased (Trampert & Fichtner, 2013). Furthermore, the existence of multiple plausible solutions that explain the data equally well, is not taken into account.

Though computationally more expensive, probabilistic (or Bayesian) source inversion, overcomes the drawbacks of the deterministic approach using a global model search that provides unbiased measures of resolution and trade-offs. Since reliable resolution analyses are essential for our study, we thus decided to apply a Bayesian approach.

4.3.1 Bayesian inverse theory

In the following we summarize Bayesian finite source inversion, using the concepts described, for instance, in Mosegaard & Tarantola (1995) and Tarantola (2005): We combine the available prior knowledge on the model parameters, the observations and the theoretical relationship between model parameters and data into the solution of the inverse problem in the form

$$\sigma(\mathbf{m}) = k \rho(\mathbf{m}) L(\mathbf{m}). \quad (4.3)$$

In equation (4.3), k is a normalization constant, \mathbf{m} denotes an element of the (26-dimensional) model space, $\rho(\mathbf{m})$ represents the prior information on the model parameters in terms of a multi-dimensional probability density function (*prior pdf*), and the likelihood function $L(\mathbf{m})$ provides a measure of how well a model, \mathbf{m} explains the data \mathbf{d} . Consequently, the solution or posterior state of information, $\sigma(\mathbf{m})$, is given as well in terms of a multi-dimensional probability density function (*posterior pdf*).

4.3.2 Prior information and likelihood function

As priors in model space we use probability densities that are constant within a prescribed interval. Furthermore, different model parameters m^α ($\alpha = 1, 2, \dots, 26$) are initially assumed to be independent. The index α refers either to slip amplitude, rise time or rupture velocity. It follows that the full prior in the model space is given by

$$\rho(\mathbf{m}) = \prod_{\alpha} \rho_{\alpha}(m^{\alpha}), \quad (4.4)$$

with

$$\rho_\alpha(m^\alpha) = \begin{cases} (m_{\max}^\alpha - m_{\min}^\alpha)^{-1}, & \text{for } m_{\min}^\alpha \leq m^\alpha \leq m_{\max}^\alpha \\ 0, & \text{otherwise.} \end{cases}$$

For the intervals $[m_{\min}^\alpha, m_{\max}^\alpha]$ we use $[0, 5]$ m for the slip amplitude, $[0.5, 1.5]$ s for the rise time and $[2000, 3000]$ m/s for the rupture velocity. Assuming negligible forward modeling errors, we employ the exponential likelihood function

$$L(\mathbf{m}) = k' \exp \left[- \sum_l \left(\frac{\chi_l(\mathbf{m})}{s_l} \right) \right], \quad (4.5)$$

where k' is again a normalization constant. In equation (4.5), $\chi_l(\mathbf{m})$ denotes the measure of misfit between data predicted from the model and the observed data, and s_l represents the noise level. The index l refers to either velocity or rotation rate recordings. We calculate the misfit based on the L_2 -norm as

$$\chi_l(\mathbf{m}) = \frac{1}{2} \sum_r (\mathbf{g}_l^r(\mathbf{m}) - \mathbf{d}_{l,\text{obs}}^r)^2. \quad (4.6)$$

The symbols $\mathbf{g}_l^r(\mathbf{m})$ and $\mathbf{d}_{l,\text{obs}}^r$ denote the forward modeling function and the (synthetic) data at station r , respectively. We underline that the noise level balances the summation of misfits computed for diverse observations so that the probabilistic inversion approach naturally accounts for different data types like velocity and rotation rate seismograms.

4.3.3 The Metropolis algorithm

To approximate the *posterior pdf*, probabilistic inversions use random walks through the model space. The specific implementation of this sampling process depends on the individual problem under investigation. As our forward problem can be solved efficiently for a large number of models, and because the dimension of the model space is small, we can apply a Metropolis-Hastings algorithm (Metropolis & Ulam, 1949; Metropolis *et al.*, 1953; Hastings, 1970). For this, we produce a candidate model, \mathbf{m}_{can} , that samples the *prior pdf*. The candidate model is then included in the final ensemble of models with an acceptance probability P_{accept} that depends on the current model \mathbf{m}_{cur}

$$P_{\text{accept}} = \begin{cases} 1, & \text{if } \chi(\mathbf{m}_{\text{can}}) \leq \chi(\mathbf{m}_{\text{cur}}) \\ \frac{L(\mathbf{m}_{\text{can}})}{L(\mathbf{m}_{\text{cur}})}, & \text{if } \chi(\mathbf{m}_{\text{can}}) > \chi(\mathbf{m}_{\text{cur}}), \end{cases} \quad (4.7)$$

with

$$\chi(\cdot) = \sum_l \left(\frac{\chi_l(\cdot)}{s_l} \right). \quad (4.8)$$

Hence the candidate model is always accepted if it improves the data fit, but if the candidate model degrades the data fit there is still a chance of not being rejected. The ensemble of accepted models can be displayed in the form of histograms, where the height of a column is proportional to the posterior likelihood

4.3.4 Shannon's measure of information gain

Once the *posterior pdf* is approximated with sufficient accuracy, we can quantify the information content or information gain relative to the prior. Following Tarantola (2005), we define the information gain for a specific inversion scenario as the relative information content of the *posterior pdf*, ρ , with respect to the *prior pdf*, σ , as

$$I(\rho; \sigma) = \int \rho(x) \log \left[\frac{\rho(x)}{\sigma(x)} \right] dx. \quad (4.9)$$

We use the logarithm base 2 in equation (4.9), so the unit of information gain is termed a *bit*.

4.4 Case studies with and without rotations

In this section we aim to recover rupture velocity, rise time and the 24 slip amplitude parameters of the target model, described in section 4.2.3, using the previously introduced probabilistic inversion technique. We study two different scenarios: In scenario I, we use the conventional 3-C velocity data of all 20 stations shown in Figure 4.1a. In scenario II, we also include rotational ground motions, but only consider 10 stations.

At this point, we emphasize two important aspects: (1) The information gain that we analyze in both scenarios is based on the same number of seismograms (66). This means that the pure quantity of data is the same. (2) The 10 stations in scenario II are randomly selected. First, for the purpose of easy illustration, we consider only one possible selection of 10 out of 20 stations (section 4.4.2). To assess the extent to which results depend on a particular selection, we then repeat the inversion for a large number of station subsets (section 4.4.3).

4.4.1 Scenario I: Inverting 3-C velocity observations at 20 stations

In scenario I, we invert 3-component velocity seismograms (3-C data) at all 20 stations. Applying the Metropolis rule from section 4.3.3 to 1 million test models, we approximate the *posterior pdf*. The 24 marginal *posterior pdf*'s for the slip amplitudes are plotted in Figure 4.5, where each panel

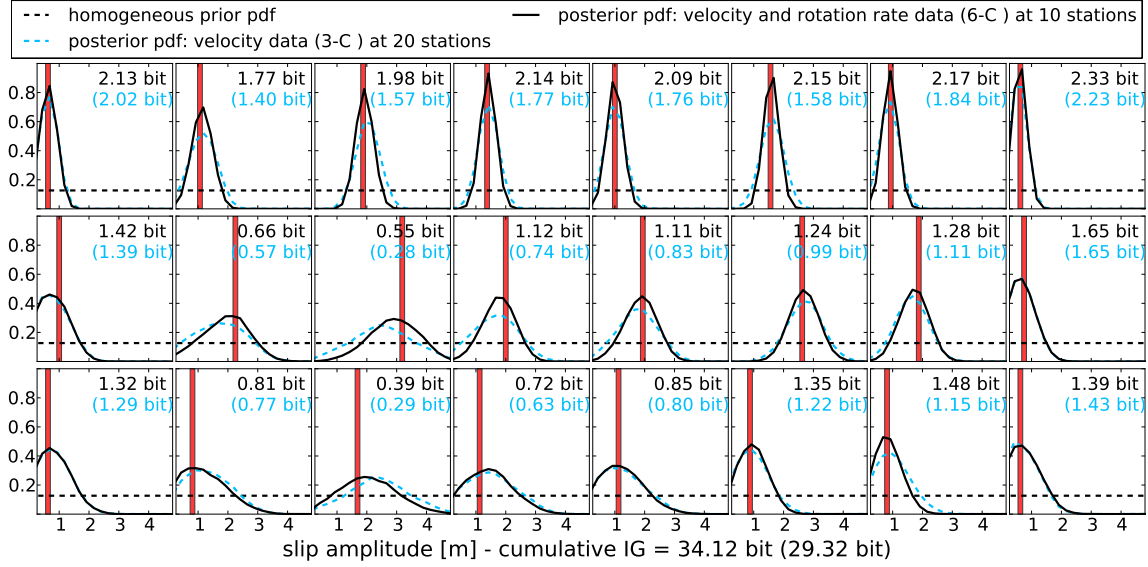


Figure 4.5: Inversion results for the slip amplitude. Each panel represents one subfault of the rupture plane arranged equivalent to Figure 4.2 while every subplot shows the *prior pdf* (black dashed line) and the true value of the target model (red bar). The blue dashed lines correspond to the *posterior pdf*'s in scenario I (3-C velocity data at 20 stations). Black lines show the same but for scenario II (6-C velocity and rotation rate data at 10 stations). The black and blue numbers describe the information gain (blue: scenario I, black: scenario II).

represents a subfault of the rupture plane, arranged as in Figure 4.2. Each subplot shows the *prior pdf* as black dashed line, and the slip value of the target model as red bar. The blue dashed lines correspond to the *posterior pdf*'s. Blue numbers in brackets describe the information gain for each parameter according to equation (4.9).

With an average of 1.77 bit, the information gain in the top row of subfaults is largest, indicating that near-surface slip amplitudes are resolved best. Their *posterior pdf*'s have well-defined peaks matching the true parameters. In the middle part of the fault plane (second row), the information gain for slip amplitudes reaches slightly more than 50% (0.95 bit on average) compared to the values observed in the top row. The same is approximately true for the bottom of the fault plane (third row).

A reliable estimate of the target slip is not feasible from several *posterior pdf*'s in the middle and the bottom rows of the fault plane. This indicates that strong motion waveforms are mostly sensitive to rupture near the surface and relatively insensitive to rupture at greater depth, in accord with previous studies (e.g. Mendoza & Hartzell, 1989; Cotton & Campillo, 1995; Semmane *et al.*, 2005; Monelli & Mai, 2008). Moreover, we attribute variations in the information content along the strike direction of the rupture plane to the irregular distribution of stations at the surface (directivity effect).

The inversion results for rupture velocity and rise time are visualized in Figure 4.6a in the form of

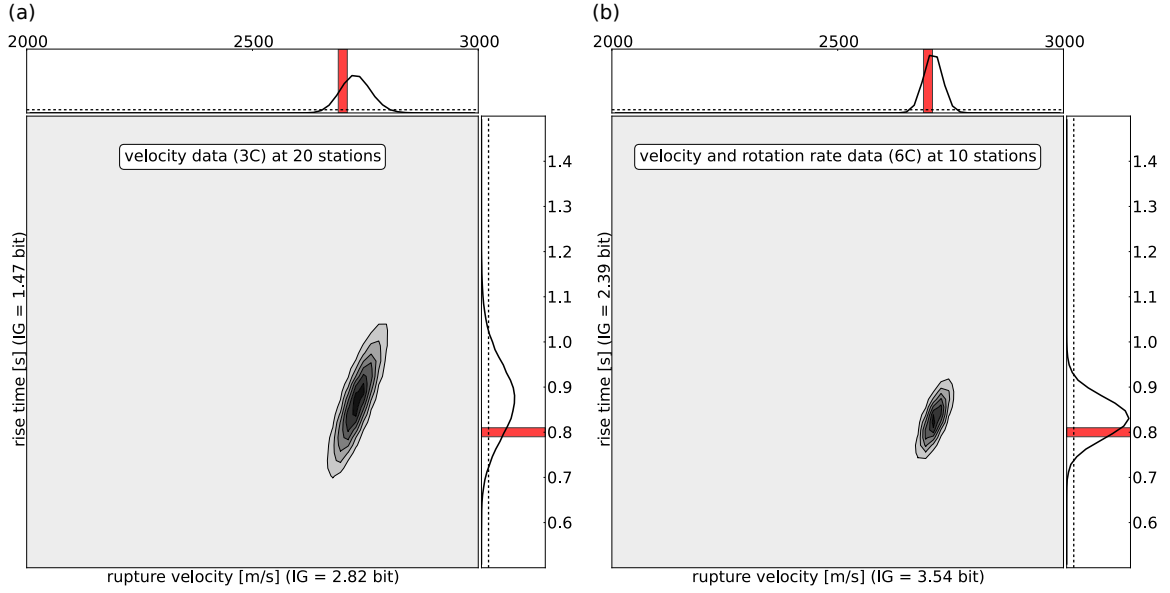


Figure 4.6: Inversion results for rupture velocity and rise time. The 2-D marginal *posterior pdf*'s express the cross-dependency between rupture velocity and rise time. The corresponding 1-D marginal *posterior pdf*'s are drawn at the top and to the right of each panel for rupture velocity and for rise time, respectively. (a) The dark gray and black shaded areas mark the relatively large region where both parameters compensate each other during the inversion process of scenario I. (b) The area of cross-dependencies between rise time and rupture time decreases substantially in scenario II.

a 2-D marginal that reveals the trade-offs between these two parameters. The dark gray and black shaded areas mark the relatively large region where variations in both parameters compensate each other, so as to have nearly no effect on the misfit. The corresponding 1-D marginal *posterior pdf*'s are drawn at the top for rupture velocity and to the right for rise time, respectively. Evidently the expected values of both approximately bell-shaped *posterior pdf*'s miss the target model, which is especially true for the rise time.

Analyzing the seismogram fit in scenario I, further motivates the inclusion of rotational ground motion observations in the following scenario II. In Figure 4.7 we compare the seismograms at station 10 for the target model (black line) and for a model randomly chosen from the 1% of models with the lowest misfit (red line). The top and the bottom row of Figure 4.7 show velocity and rotation rate seismograms, respectively. Each component is characterized by the normalized L_2 -misfit between synthetics and observations. As in the inversion process of scenario I only velocity seismograms are fitted, we observe a substantially higher L_2 -misfit in rotation rate seismograms compared to velocity seismograms. The L_2 -misfit of the vertical rotation rate seismograms is, for example, twice as large as the misfit in both horizontal velocity seismograms. This indicates the potential of inverting the source parameters including rotation rate seismograms. Figure 4.7

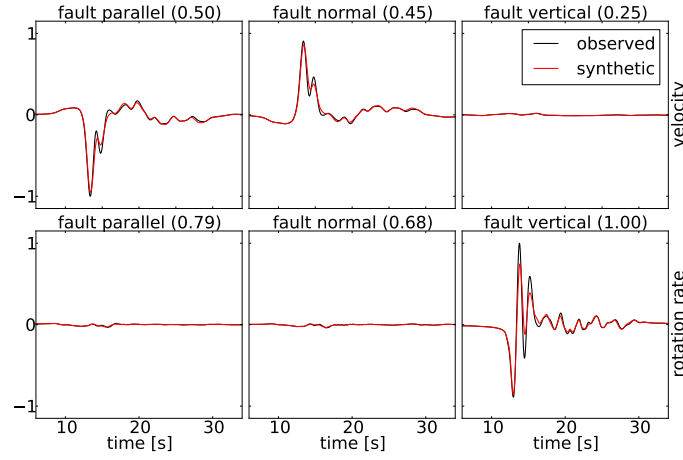


Figure 4.7: Normalized seismograms at station 10 produced by a model selected from the 1% best data fitting model samples. The numbers in brackets represent the normalized L_2 -misfits between observed (black line) and synthetic (red line) seismograms. The maximal amplitude for non normalized velocity and rotation rate seismograms is 10.1 cm/s and 3.1×10^{-5} rad/s, respectively

suggests that additional contributions from rotational ground motions may come mainly from their amplitudes, as the phases of observed and synthetic seismograms are already in good agreement.

4.4.2 Scenario II: Inverting 6-C observations at 10 stations

In contrast to scenario I, scenario II combines velocity and rotation rate observations. To keep the total amount of data in the inversion constant, we randomly select 10 stations from the complete array. These are marked as red triangles in Figure 4.1a. Despite considering 6-C data at 10 stations instead of 3-C data at 20 stations, the inversion is performed identically to scenario I.

We obtain the following results: The *posterior pdf*'s of the slip amplitude (solid black lines in Figure 4.5) in the top row subfaults acquire sharper peaks, accurately matching the target model values. Analogously to scenario I, the information gain decreases with depth. Nevertheless, 6-C observations compared to 3-C observations provide a higher information gain for slip amplitude for almost all subfaults, despite using only half as many stations. In numbers, the cumulative information gain for slip amplitude is 34.12 bit in scenario II. In scenario I, this value equals 29.32 bit, i.e. 86% of what is obtained in scenario II.

Using 6-C data at 10 stations significantly sharpens the *posterior pdf*'s for rise time and rupture velocity, as shown in the 2-D marginal of Figure 4.6. Trade-offs between the parameters are strongly reduced, and the peaks of the 1-D marginal *posterior pdf*'s are close to the target values. Compared to scenario I, the information gain in scenario II is 25 % larger for rupture velocity and more than 60 % larger for rise time.

4.4.3 Generalization: Random selection of stations

The inversion of source parameters in scenario II suggests that including rotational ground motion measurements reduces the non-uniqueness of finite source inversions. To ensure that information gain does not result from merely increasing the quantity of data, we only used 10 stations for the 6-C inversion, thus keeping the same total number of seismograms as in the 3-C scenario with 20 stations.

Repeating the inversion procedure for 1000 randomly selected combinations of 10 stations, allows us to assess the generality of our results. Figure 4.8 displays the information gain for slip amplitude, rupture velocity and rise time for all of the 1000 probabilistic inversions based on 6-C data. The information gain for the 3-C inversion with 20 stations is shown for comparison. Using 6-C data at 10 stations almost always provides more information than using 3-C data at 20 stations, regardless of which combination of 10 stations is used. This effect is most pronounced for slip amplitude in the top row of subfaults, rupture velocity, and rise time.

Only for some station combinations the information gain is lower, probably because the randomly selected stations may lead to large azimuthal gaps in the station distribution. Also, the random selection may contain little energy in horizontal rotation rate components, that we consider essential for reducing the non-uniqueness. As seen in the second and the third panel of Figure 4.8 this strongly affects the less resolved slip amplitudes in the middle and the bottom row of the fault plane. The histograms provide several peaks in contrast to the well resolved parameters indicating that the station distribution is critical for slip at greater depth.

At this point, we conclude that the non-uniqueness in finite source inversions can be significantly reduced by including rotational ground motion measurements. We emphasize that only the combination of both data types, velocity and rotation rate seismograms, leads to the benefits described above. Inverting rotation rate seismograms separately for all stations resembled the results corresponding to scenario I.

4.4.4 Requirements on the measurement accuracy of rotational ground motion sensors

It remains to be investigated which level of measurement accuracy rotational ground motion sensors must provide so that finite source inversion can benefit from the novel data type.

To address this question, we consider again the setup of scenario II, and we perform the same inversion process as described in section 4.4.2 with varying noise levels in the rotation rate data. Since the information gain of 6-C relative to 3-C inversions is controlled by the ratio between velocity and rotation rate noise levels, we keep the noise level in velocity seismograms constant at 1% of the maximum amplitude of all velocity seismograms.

In a first scenario, we contaminate the rotation rate synthetics with Gaussian noise that has an amplitude 0.5% of the maximum rotation rate amplitude at all stations, i.e. half the noise amplitude

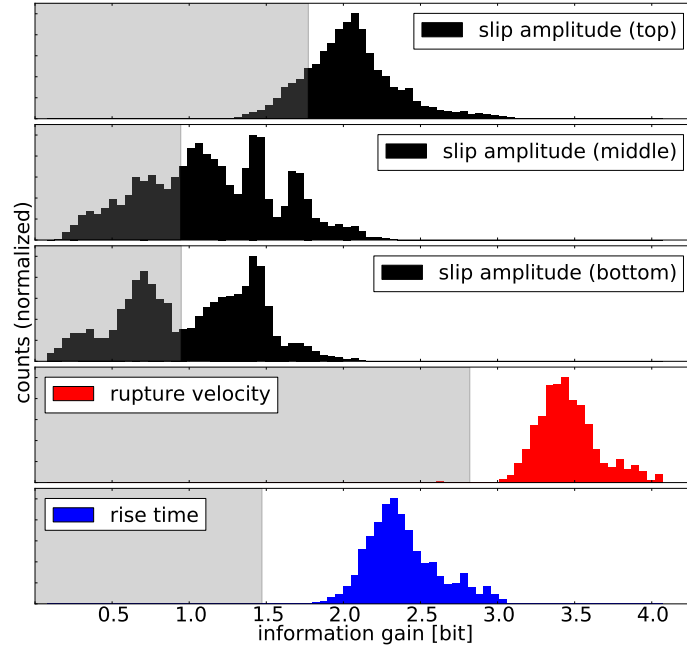


Figure 4.8: Histograms for the information gain corresponding to 1000 scenarios inverting 6-C observations at 10 randomly selected stations. Black histograms correspond to the average information gain of slip amplitudes for subfaults in the first (top), second (middle) and third (bottom) row of the fault plane. Histograms for rupture velocity and rise time are plotted in red and blue, respectively. For comparison the information gain of scenario I is indicated by a gray bar in each panel.

of scenario II. As illustrated in Figure 4.9, this results in a cumulative information gain for all parameters of 45 bit, compared to 33 bit in scenario I. The cumulative information gain of scenario II (1% noise) equals 40.48 bit. Considering various scenarios with increasing noise levels in the rotation rates provides the black curve in Figure 4.9. As expected, increasing the noise level in the rotation rates decreases the cumulative information gain. When the rotation rate noise level reaches 1.75%, the cumulative information gain drops to the one of scenario I, meaning that the potentially available additional information is full compensated by measurement errors. For a noise level of 5% the cumulative information gain is already lower than 30 bit.

In summary, our test with varying noise levels suggests that the signal-to-noise ratio in rotation rate recordings should be similar to the one in velocity recordings in order to be beneficial for finite source inversions.

4.5 Discussion

The main result of this study is that the incorporation of rotational ground motion data into kinematic source inversion can significantly reduce non-uniqueness, provided that measurement errors in rotation data are roughly at the same level as errors in classical translation data. This result must

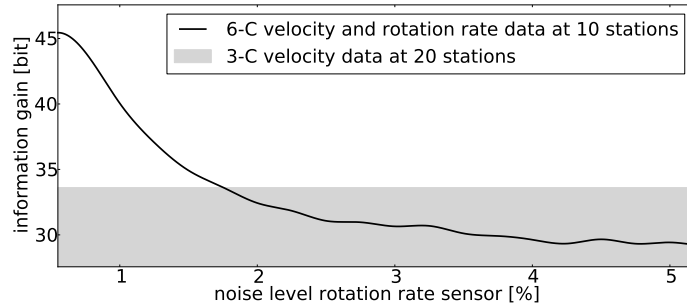


Figure 4.9: Cumulative information gain corresponding to different noise levels in the rotation rate sensors and constant noise level in the velocity observations. For comparison the cumulative information gain received in scenario I, that inverts only velocity seismograms, is indicated by the gray bar.

be seen in the light of the following aspects that we discuss below: (1) The availability of current and the construction of future rotation sensors, (2) the physical origin of the observed information gain, (3) the potential implications for earthquake physics and ground motion predictions, and (4) the inversion technique used in this study.

4.5.1 Measuring rotational ground motions

While the vertical component of rotational ground motion can be approximated via finite differences from translational recordings at the surface, a similar approximation of the horizontal components would require borehole seismometers that are mostly unavailable. Alternatively, the seismogeodetic method (Bodin *et al.*, 1997; Spudich *et al.*, 1995; Spudich & Fletcher, 2008) may be used to derive rotational ground motion from translation recordings at the surface (e.g. Suryanto *et al.*, 2006; Kendall *et al.*, 2012; Pham *et al.*, 2012). This method requires, however, detailed knowledge of 3-D subsurface structure. In any case, the derivation of rotations from translations relies on the availability of a sufficiently dense array where the inter-station spacing is shorter than the minimum wavelength. This condition would not be met in our study where the shortest wavelength is around 3 km.

A high-resolution, field-deployable and affordable instrument that measures rotation is currently not available. However, the construction of rotation sensors based on ring-laser technology (e.g. Schreiber *et al.*, 2009; Velikoseltsev *et al.*, 2012), adapted gyroscopes (e.g. Bernauer *et al.*, 2012), or systems of multiple geophones (e.g. Brokešová & Málek, 2010; Brokešová *et al.*, 2012) is rapidly progressing. In the course of these instrument developments, the admissible measurement errors for specific applications must be established. To be beneficial in finite source studies, rotation sensors should reach approximately the same signal-to-noise ratio as classically used seismometers that measure translational ground motion.

4.5.2 Physical origin of the information gain

Probabilistic inversion does not directly provide a physical explanation for an observed information gain. It seems plausible, however, that the horizontal rotation components make a significant contribution. Horizontal ground rotations carry information on the vertical displacement gradient, thereby providing information that cannot be obtained from conventional 3-C recordings on the Earth's surface. It follows that a potentially positive information gain from the incorporation of rotational ground motions may be even more significant for events that fracture as well in the vertical direction (dip-slip), thereby producing larger horizontal rotation.

4.5.3 Implications for earthquake physics and ground motion predictions

Our analysis shows that the incorporation of rotational ground motions is most beneficial for the resolution of rupture velocity and rise time, i.e. those parameters that are non-linearly related to ground motion (Archuleta, 1984). Accurate knowledge of both rupture velocity and rise time is critical for the inference of rheological fault properties and friction law parameters, i.e. the earthquake dynamics (e.g. Tinti *et al.*, 2009).

Rupture velocity controls the directivity effect which increases ground motions normal to the fault (e.g. Somerville *et al.*, 1997). Furthermore, far field displacement depends on the moment time derivative, meaning that ground motions depend strongly on the rise time. In this context, the incorporation of rotational ground motions has the potential to improve estimates of parameters that are critical for earthquake-induced ground motion.

4.5.4 Probabilistic source inversion

Probabilistic inversion bring numerous advantages that are critical for our study, including the absence of linearization and regularization, the nearly exhaustive exploration of the model space, and the possibility to obtain comprehensive information on uncertainties and trade-offs. Critical aspects of probabilistic inversion are the choice of a prior and the convergence of the sampling algorithm.

The incorporation of prior knowledge is a controversially debated component of Bayesian inversion (e.g. Efron, 2013). It must, most importantly, be acknowledged that any inference is relative to the prior, and this prior must be taken into account when interpreting the results. In our case it is reasonable to model the available prior knowledge in terms of a flat *prior pdf* that reflects a low level of knowledge. The impact of this choice is, however, small because we are primarily interested in the comparison of the information gain of various scenarios with identical priors.

To assess the convergence of the Metropolis-Hastings algorithm, we monitor the change of the marginal *posterior pdf*'s as a function of the total number of samples. This is displayed in figure 4.10 for scenario II (section 4.4.2). After approximately 250,000 samples, i.e. 25% of the total one million samples, the observed changes in the *posterior pdf*'s are practically negligible. This

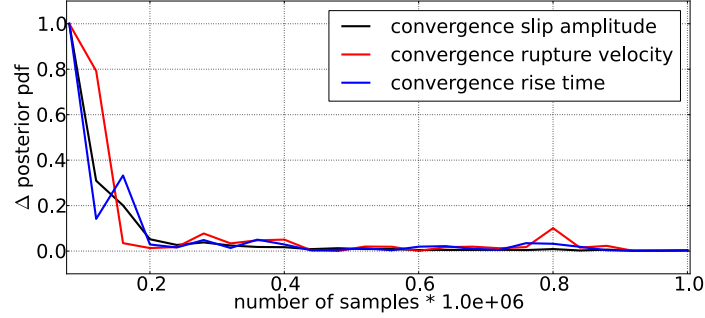


Figure 4.10: Convergence of the Metropolis-Hastings sampling algorithm. The change of the marginal *posterior pdf*'s for slip amplitude (black line), the rupture velocity (red line) and the rise time (blue line) are plotted as a function of accumulated samples.

indicates that the sampling process indeed converges to the *posterior pdf*. In scenario I, we observed a comparable convergence behavior. Furthermore, we note that various independently repeated runs resembled the *posterior pdf*'s presented in both scenarios. Thus we can assume that our results are not affected by the starting model of the sampling process.

4.6 Conclusions

Based on a series of Bayesian finite source inversions, we assess the potential of rotational ground motion recordings to reduce non-uniqueness in kinematic source parameters. Our results suggest that the incorporation of rotational ground motion recordings can significantly reduce non-uniqueness in finite source inversions when measurement uncertainties are approximately similar to or below the uncertainties of translational velocity recordings. This requirement should be met by currently developed rotation sensors in order to be useful in finite source studies. Rupture velocity and rise time have the highest potential to benefit from rotation data. The trade-offs between both parameters can be strongly reduced, and the information gain significantly increased. It follows that rotation data can improve secondary inferences that depend on rupture velocity and rise time, such as radiation directivity and frictional properties of the fault. The reduced non-uniqueness most likely results from the horizontal rotation components that carry information on the vertical displacement velocity gradient that cannot be derived from translation recordings at the surface.

Acknowledgments

The authors would like to thank Amir Khan for appreciated considerations about Bayesian inverse theory, Lion Krischer for the support in professional code optimisation, and Fabrice Cotton for useful discussions of our results. Furthermore, we thank the members of the seismology groups at LMU Munich, Universiteit Utrecht and ETH Zurich for valuable discussions. The research presented in this article was supported by the International Graduate School THESIS within the

Bavarian Elite Network, the European Commission (Marie Curie Actions, ITN QUEST) and the German Research Foundation (IG 16/15-1). The numerical computations were performed also on the National Supercomputer SuperMUC maintained by the Leibniz-Rechenzentrum.

Appendix

A.1 Details for the computation of sensitivity kernels

The forward wavefield is excited by a bandpass filtered Heaviside function between 20 s and 200 s. The moment tensor components given in Nm are

$$\begin{aligned}M_{\theta\theta} &= 0.710 \cdot 10^{19} \\M_{\phi\phi} &= -0.356 \cdot 10^{19} \\M_{zz} &= -0.355 \cdot 10^{19} \\M_{\theta\phi} &= 0.800 \cdot 10^{19} \\M_{\theta z} &= 0.315 \cdot 10^{19} \\M_{\phi z} &= -1.150 \cdot 10^{19}.\end{aligned}$$

The adjoint sources for the sensitivity kernels $K_m(w)$ and $K_m(s)$ are dipolar sources described by the moment tensor \mathbf{M} . The explicit moment tensor components for the computation of $K_m(w)$ are

$$\begin{aligned}M_{\theta\theta} &= M_{\phi\phi} = M_{zz} = 0 \\M_{\theta\phi} &= \frac{1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_z(\mathbf{x}^r, t) \\M_{\theta z} &= \frac{-1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_\phi(\mathbf{x}^r, t) \\M_{\phi\theta} &= \frac{-1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_z(\mathbf{x}^r, t) \\M_{\phi z} &= \frac{1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_\theta(\mathbf{x}^r, t) \\M_{z\theta} &= \frac{1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_\phi(\mathbf{x}^r, t) \\M_{z\phi} &= \frac{-1}{2 \int \omega^2(\mathbf{x}^r, t) dt} \omega_\theta(\mathbf{x}^r, t).\end{aligned}$$

The moment tensor components corresponding to $K_m(s)$ are given by

$$\begin{aligned}
 M_{\theta\theta} &= \frac{-1}{\int [\text{tr } \mathbf{e}(\mathbf{x}^r, t)]^2 dt} \text{tr } \mathbf{e}(\mathbf{x}^r, t) \\
 M_{\phi\phi} &= \frac{-1}{\int [\text{tr } \mathbf{e}(\mathbf{x}^r, t)]^2 dt} \text{tr } \mathbf{e}(\mathbf{x}^r, t) \\
 M_{zz} &= \frac{-1}{\int [\text{tr } \mathbf{e}(\mathbf{x}^r, t)]^2 dt} \text{tr } \mathbf{e}(\mathbf{x}^r, t) \\
 M_{\theta\phi} &= M_{\theta z} = M_{\phi\theta} = M_{\phi z} = M_{z\theta} = M_{z\phi} = 0.
 \end{aligned}$$

The sensitivity kernel $K_m(v)$ is computed via the adjoint source time function

$$f_k^v(\mathbf{x}) = \frac{1}{\int \mathbf{v}^2(\mathbf{x}^r, t) dt} \ddot{u}_k(\mathbf{x}^r) \delta(\mathbf{x} - \mathbf{x}^r).$$

For a detailed derivation of the adjoint source time functions we refer to Fichtner & Igel (2009).

B.1 Optimal eigenvalue

Given $\sum_{i=1}^n w_i^2 = 1$ and $\frac{\partial \mathcal{L}}{\partial w_l} = 0$ we have $\mathcal{L}(\mathbf{w}, \lambda) = \lambda$.

Proof:

$$\begin{aligned}
 \mathcal{L}(\mathbf{w}, \lambda) &= \sum_{j=1}^p b_j \int_V \left[\sum_{i=1}^n w_i K_{ij}(\mathbf{x}) \right]^2 d^3 \mathbf{x} - \lambda \left(\sum_{i=1}^n w_i^2 - 1 \right) \\
 &= \sum_{j=1}^p b_j \int_V \left[\sum_{i=1}^n w_i K_{ij}(\mathbf{x}) \right]^2 d^3 \mathbf{x} \\
 &= \sum_{j=1}^p b_j \int_V [w_1 w_1 K_{1j} K_{1j} + \dots + w_1 w_n K_{1j} K_{nj} + \\
 &\quad \vdots \quad \quad \quad \vdots \\
 &\quad + w_n w_1 K_{nj} K_{1j} + \dots + w_n w_n K_{nj} K_{nj}] d^3 \mathbf{x} \\
 &= w_1 M_{11} w_1 + \dots + w_1 M_{1n} w_n + \\
 &\quad \vdots \quad \quad \quad \vdots \\
 &\quad + w_n M_{n1} w_1 + \dots + w_n M_{nn} w_n \\
 &= \sum_{l,i=1}^n w_l M_{li} w_i = \mathbf{w}^T \mathbf{M} \mathbf{w} = \lambda \mathbf{w}^T \mathbf{w} = \lambda
 \end{aligned}$$

B.2 Sensitivity kernels for vertical density gradients

To derive expressions for sensitivity kernels with respect to vertical density gradients, $\partial_r \rho$, we start with the general first-order expression that relates a change in the data δd to a change in density, $\delta \rho$:

$$\delta d = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^R K_{\rho}(\theta, \phi, r) \delta \rho(\theta, \phi, z) d\theta d\phi dr, \quad (\text{B.1})$$

where R denotes the radius of the Earth. In equation (B.1), we incorporated the term $r^2 \sin \theta$ in the spherical volume element in the definition of the density kernel K_{ρ} , thereby leaving a simple integral of $K_{\rho} \delta \rho$ over colatitude θ , longitude ϕ and radius r . Substituting the identity

$$K_{\rho}(\theta, \phi, r) = \partial_r \int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr' \quad (\text{B.2})$$

into equation (B.1), yields

$$\delta d = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^R \left[\partial_r \int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr' \right] \delta \rho(\theta, \phi, z) d\theta d\phi dr. \quad (\text{B.3})$$

Integrating by parts with respect to r , transforms (B.3) into

$$\begin{aligned} \delta d = & \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \left[\int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr' \right] \delta \rho(\theta, \phi, z) d\theta d\phi \Big|_{r=0}^R \\ & - \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^R \left[\int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr' \right] \partial_r \delta \rho(\theta, \phi, z) d\theta d\phi dr. \end{aligned} \quad (\text{B.4})$$

Recognising that the vertical integral over the density kernel K_{ρ} in $\alpha - \beta - \rho$ parametrisation is nearly zero (e.g. Takeuchi & Saito, 1972), the first term on the right-hand side of (B.4) can be ignored. Using the identity $\partial_r \delta \rho = \partial_r (\rho - \rho_0) = \partial_r \rho - \partial_r \rho_0 = \delta \partial_r \rho$, equation (B.4) simplifies to

$$\delta d = - \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^R \left[\int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr' \right] \delta \partial_r \rho(\theta, \phi, z) d\theta d\phi dr. \quad (\text{B.5})$$

From equation (B.5) we see that the sensitivity kernel with respect to the vertical density gradient $\partial_r \rho$ is given by the negative integral of K_{ρ} with respect to r :

$$K_{\partial_r \rho}(\theta, \phi, r) = - \int_{r'=0}^r K_{\rho}(\theta, \phi, r') dr'. \quad (\text{B.6})$$

With $K_{\partial_r \rho}$ we retrieve the generic expression relating a change in data to a change in model parameters:

$$\delta d = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^R K_{\partial_r \rho}(\theta, \phi, r) \delta \partial_r \rho(\theta, \phi, z) d\theta d\phi dr. \quad (\text{B.7})$$

Kernels for relative perturbations $\delta \ln \partial_r \rho$ are obtained from $K_{\partial_r \rho}$ by multiplication with a reference value for $\partial_r \rho$. In the case of the 1-D reference Earth model ak135 (Kennett *et al.*, 1995), used in the examples of section 3.3, $\partial_r \rho$ takes the nearly constant value of $6.05 \cdot 10^{-4} \text{ kg/m}^4$ from 35-410 km depth.

C.1 1-D Earth model

For the computation of the seismic wave field for the moment tensor point sources we use the 1-D Earth model from Semmane *et al.* (2005):

Depth [km]	V_p [km/s]	V_s [km/s]	ρ [kg/m ³]
0	5.50	3.179	2600
2	6.05	3.497	2700
16	6.60	3.815	2800
38	8.03	4.624	3100

C.2 Source function

In this study the source function is constructed as an ordinary ramp function. In the time domain and as a function of the rise time R the ramp function is expressed as

$$\tilde{S}(R, t) = \begin{cases} 0, & t < 0 \\ t/R, & 0 \leq t \leq R \\ 1, & t > R. \end{cases}$$

The representation in the frequency domain can then be computed as

$$S(R, \omega) = \mathcal{F}\left[\frac{d}{dt}\tilde{S}\right](\omega) \frac{1}{i\omega} = \frac{1}{\omega^2 R} [\exp(-i\omega R) - 1]$$

where \mathcal{F} denotes the Fourier transform of $\frac{d}{dt}\tilde{S}$.

C.3 Analyzing the results of scenario II from section 4.4.2

Highlighting some characteristic features of probabilistic inversion, this section aims to evaluate the ensemble of models accepted during the inversion process of scenario II from section 4.4.2.

To acquire intuition for the variability of plausible models, 30 of the 10,000 best data fitting models were randomly selected. These are displayed in Figure C.1 together with their respective deviations from the target model. While the data misfits of most models fall within the narrow range of 4.2 ± 0.2 , significant variability exists for those subfaults where the information gain is comparatively low, i.e. at greater depth and near the north-western tip of the fault plane.

In Figure C.2 each column corresponds to one component of the 6-C observations and each row represents one of the 10 stations randomly selected in scenario II. The normalized scaling factor encodes the strength of the signal at each station. While decreasing scaling factors correspond to smaller amplitudes the highest amplitudes are observed in station 19. Figure C.2 indicates,

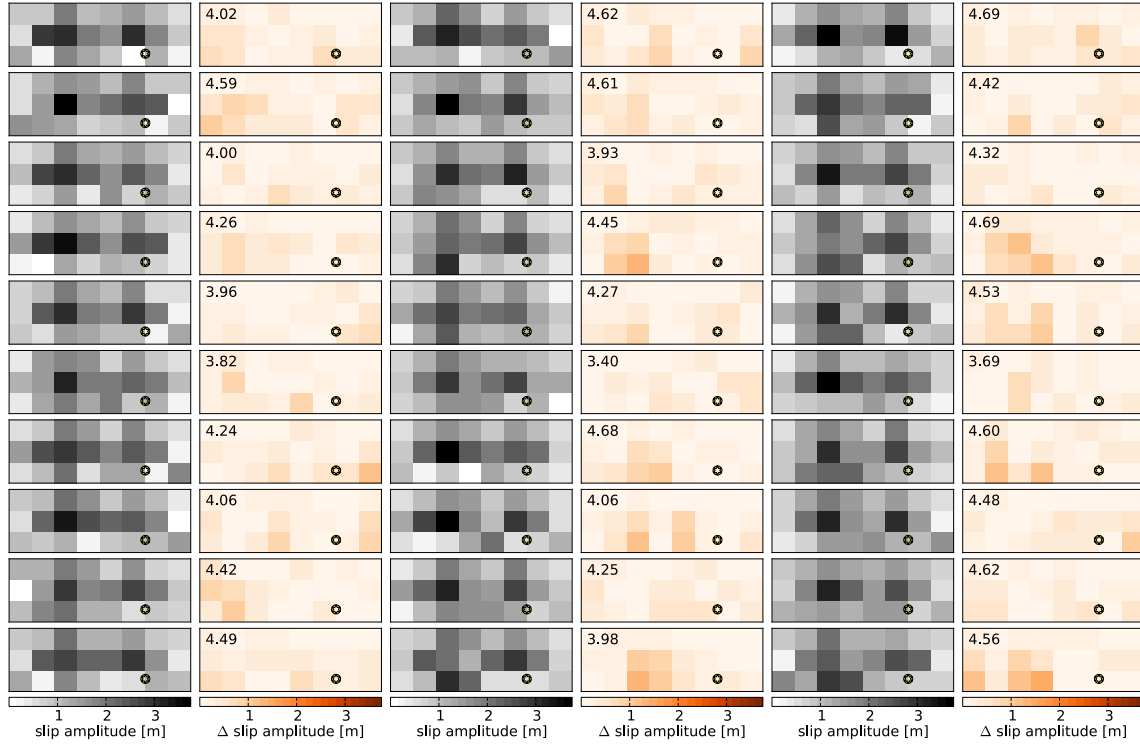


Figure C.1: 30 slip amplitude models (1st, 3rd and 5th column) randomly picked from the 10,000 best data fitting models in the sampled ensemble. The 2nd, 4th and 6th column quantify the corresponding discrepancies between the slip amplitudes of each subfault compared to the target model. The numbers indicate the L_2 -misfit between the individual models and the target model depending on slip amplitudes, rupture velocity and rise time.

as expected, that seismograms with high energy at stations close to the fault, e.g. station 4, 5, 14 and 19 (see Figure 4.3) are fitted best while slightly larger variations are observed for stations recording lower amplitudes (e.g. station 1, 2 and 11). The fact that different models produce similar seismograms despite considerable variations concerning individual parameters expresses the non-uniqueness of the inverse problem and reflects the trade-offs between the model parameters and the observations. For example the slip amplitudes of the 1st model plotted in column 5 of Figure C.1 resemble the target model while, despite of a smaller L_2 -misfit value, the slip amplitudes of the last model in column 3 miss the target model concerning the slip amplitudes of several subfaults. This is possible because the misfit depends also on the values for the rupture velocity and the rise time and it demonstrates the problem of providing a single model as the solution of probabilistic inverse scenarios.

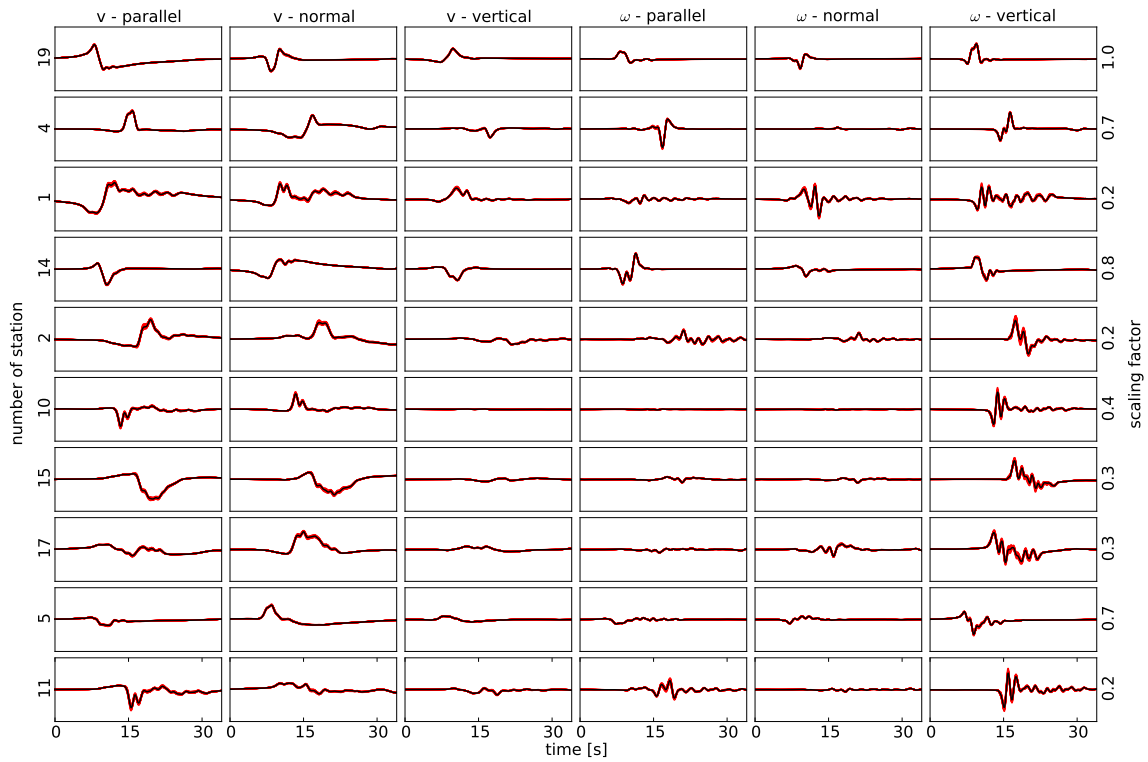


Figure C.2: From the left to the right: fault parallel, normal and vertical velocity followed by rotation rate seismograms in the same order. In each panel black lines represent observations while red lines correspond to the simulated seismograms of all 30 models from Figure C.1. Each row corresponds to one of the 10 stations selected in scenario II. The normalized scaling factor to the right encodes the strength of the signal at each station. Note that a decreasing scaling factor correspond to smaller amplitudes. According to that the highest amplitudes are observed at station 19.

Bibliography

- Agnew, D. C. & Wyatt, F. K., 2003. Long-base laser strainmeters: A review. Technical report, Scripps Institution of Oceanography, University of California, San Diego. <http://escholarship.org/uc/item/21z72167> [Last accessed on 2014-01-11].
- Aki, K. & Richards, P., 2002. *Quantitative Seismology*. University Science Books.
- Archuleta, R. J., 1984. A faulting model for the 1979 Imperial Valley earthquake. *J. Geophys. Res.*, **89**, 4559–4585.
- Babuška, V. & Cara, M., 1991. *Seismic anisotropy in the Earth*. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Backus, G. E., 1962. Long-wave elastic anisotropy produced by horizontal layering. *J. Geophys. Res.*, **67**, 4427–4440.
- Backus, G. E. & Gilbert, F., 1967. Numerical application of a formalism for geophysical inverse problems. *Geophys. J. Roy. Astr. Soc.*, **13**, 247–276.
- Backus, G. E. & Gilbert, F., 1968. The resolving power of gross Earth data. *Geophys. J. Roy. Astr. Soc.*, **16**, 169–205.
- Benioff, H., Gutenberg, B. & Richter, C. F., 1951. Progress Report, California Institute of Technology, 1950. *Trans. Am. Geophys. Union*, **32**, 749–754.
- Beresnev, I. A., 2003. Uncertainties in Finite-Fault Slip Inversions: To What Extent to Believe? (A Critical Review). *Bull. Seismol. Soc. Am.*, **93**, 2445–2458.
- Bernauer, F., Wassermann, J. & Igel, H., 2012. Rotational sensors - a comparison of different sensor types. *J. Seismol.*, **16**, doi: 10.1007/s10950-012-9286-7.
- Bernauer, M., Fichtner, A. & Igel, H., 2009. Inferring earth structure from combined measurements of rotational and translational ground motions. *Geophysics*, **74**, WCD41–WCD47.
- Bernauer, M., Fichtner, A. & Igel, H., 2012. Measurements of translation, rotation and strain: new approaches to seismic processing and inversion. *J. Seismol.*, **16**, 669–681.

- Bilham, R., 2009. The seismic future of cities. *Bulletin of Earthquake Engineering*, doi: 10.1007/s10518-009-9147-0.
- Blum, J., Igel, H. & Zumberge, M., 2010. Observations of rayleigh-wave phase velocity and coseismic deformation using an optical fiber, interferometric vertical strainmeter at the safod borehole, california. *Bull. Seismol. Soc. Am.*, **100**, doi: 10.1785/0120090333.
- Bodin, P., Gomberg, J., Singh, S. K. & Santoyo, M., 1997. Dynamic deformations of shallow sediments in the Valley of Mexico, Part i: Three-dimensional strains and rotations recorded on a seismic array. *Bull. Seismol. Soc. Am.*, **87**, 528–539.
- Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm. *Geophys. J. Int.*, **178**, 1411–1436.
- Brokešová, J. & Málek, J., 2010. New portable sensor system for rotational seismic motion measurements. *Rev. Sci. Instrum.*, **81(8):084501**, doi: 10.1063/1.3463271.
- Brokešová, J., Málek, J. & Kolínský, P., 2012. Rotaphone, a mechanical seismic sensor system for field rotation rate measurements and its in situ calibration. *J. Seis.*, **16**, 603–621.
- Capdeville, Y., Guillot, L. & Marigo, J. J., 2010. 2-D nonperiodic homogenization to upscale elastic media for P-SV waves. *Geophys. J. Int.*, **182**, 903–922.
- Cara, M., Lévêque, J. J. & Maupin, V., 1984. Density-versus-depth models from multimode surface waves. *Geophys. Res. Lett.*, **11**, 633–636.
- Chaljub, E., Komatitsch, D., Vilotte, J., Capdeville, Y., Valette, B. & Festa, G., 2007. Spectral-Element Analysis in Seismology. *Adv. Geophys.*, **48**, 365–419.
- Chen, P., 2011. Full-wave seismic data assimilation: Theoretical background and recent advances. *Geophys. J. Int.*, **168**, 1527–1552.
- Chou, C. W. & Booker, J. R., 1979. A Backus-Gilbert approach to inversion of travel-time data for three-dimensional velocity structure. *Geophys. J. R. astr. Soc.*, 325–344.
- Cochard, A., Igel, H., Schuberth, B., Suryanto, W., Velikoseltsev, A., Schreiber, U., Wassermann, J., Scherbaum, F. & Vollmer, D., 2006. Rotational motions in seismology: theory, observation, simulation. In R. Teisseyre, M. Takeo, & E. Majewski (Eds.), *Earthquake Source Asymmetry, Structural Media and Rotation Effects*, pp. 391–411. Springer, Heidelberg.
- Cotton, F. & Campillo, M., 1994. Application of seismogram synthesis to the study of earthquake source from strong motion records. *Annali di Geofisica*, **XXXVII(6)**, 1539–1564.
- Cotton, F. & Campillo, M., 1995. Frequency domain inversion of strong motions: Application to the 1992 Landers earthquake. *J. Geophys. Res.*, **100**, doi: 10.1029/94JB0212.

- Curtis, A. & Maurer, H., 2000. Optimizing the design of geophysical experiments: Is it worthwhile? *The Leading Edge*, **19**, 1058–1062.
- Custodio, S., Liu, P. & Archuleta, R. J., 2005. The 2004 M_w 6.0 parkfield, california, earthquake: inversion of near-source ground motion using multiple data sets. *Geophys. Res. Lett.*, **32**, doi: 10.1029/2005GL024417.
- Dahlen, F. A. & Baig, F. A., 2002. Fréchet kernels for body wave amplitudes. *Geophys. J. Int.*, **150**, 440–466.
- Dalton, C. A., Ekström, G. & Dziewonski, A. M., 2008. The global attenuation structure of the upper mantle. *J. Geophys. Res.*, **113**, doi:10.1029/2007JB005429.
- Debayle, E. & Ricard, Y., 2013. Seismic observations of large-scale deformation at the bottom of fast-moving plates. *Earth Planet. Sci. Lett.*, in press.
- Dziewoński, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model. *Phys. Earth Planet. Inter.*, **25**, 297–356.
- Efron, B., 2013. Bayes' Theorem in the 21st Century. *Science*, **340**, doi:10.1126/science.1236536.
- Ferreira, A. M. G. & Igel, H., 2009. Rotational motions of seismic surface waves in a laterally heterogeneous earth. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080149.
- Ferreira, A. M. G., Woodhouse, J. H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography. *J. Geophys. Res.*, **115**, doi:1029/2009JB006716.
- Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion*. Springer, Heidelberg.
- Fichtner, A., Bunge, H.-P. & Igel, H., 2006a. The adjoint method in seismology - I. Theory. *Phys. Earth Planet. Inter.*, **157**, 86–104.
- Fichtner, A., Bunge, H.-P. & Igel, H., 2006b. The adjoint method in seismology - II. Applications: traveltimes and sensitivity functionals. *Phys. Earth Planet. Inter.*, **157**, 105–123.
- Fichtner, A. & Igel, H., 2008. Efficient numerical surface wave propagation through the optimization of discrete crustal models - a technique based on non-linear dispersion curve matching (DCM). *Geophys. J. Int.*, **173**, 519–533.
- Fichtner, A. & Igel, H., 2009. Sensitivity densities for rotational ground motion measurements. *Bull. Seis. Soc. Am.*, **99**, 1302–1314.
- Fichtner, A., Kennett, B. L. N., Igel, H. & Bunge, H.-P., 2009. Spectral-element simulation and inversion of seismic waves in a spherical section of the Earth. *J. Num. An. Ind. Appl. Math.*, **4**, 11–22.

- Fichtner, A., Kennett, B. L. N., Igel, H. & Bunge, H.-P., 2010. Full waveform tomography for radially anisotropic structure: New insight into present and past states of the Australasian upper mantle. *Earth Planet. Sci. Lett.*, **290**, 270–280.
- Fichtner, A., Kennett, B. L. N. & Trampert, J., 2012. Separating intrinsic and apparent anisotropy. *Phys. Earth Planet. Int.*, **219**, 11–20.
- Fichtner, A., Saygin, E., Taymaz, T., Cupillard, P., Capdeville, Y. & Trampert, J., 2013. The deep structure of the North Anatolian Fault Zone. *Earth Planet. Sci. Lett.*, **373**, 109–117.
- Fichtner, A. & Tkalcic, H., 2010. Insights into the kinematics of a volcanic caldera drop: Probabilistic finite-source inversion of the 1996 Bardarbunga, Iceland, earthquake. *Earth Planet. Sci. Lett.*, **297**, 607–615.
- Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion. *Geophys. J. Int.*, **187**, 1604–1624.
- Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y. & Villasenor, A., 2013. Multi-scale full waveform inversion. *Geophys. J. Int.*, doi: 10.1093/gji/ggt118.
- Fishwick, S., Kennett, B. L. N. & Reading, A. M., 2005. Contrasts in lithospheric structure within the Australian Craton. *Earth Planet. Sci. Lett.*, **231**, 163–176.
- Fukuyama, E., Ellsworth, W. L., Waldhauser, F. & Kubo, A., 2003. Detailed Fault Structure of the 2000 Western Tottori, Japan, Earthquake Sequence. *Bull. Seismol. Soc. Am.*, **93**, 1468–1478.
- Gomberg, J. & Agnew, D., 1996. The accuracy of seismic estimates of dynamic strains: An evaluation using strainmeter and seismometer data from Piñon Flat observatory, California. *Bull. Seismol. Soc. Am.*, **86**, 212–220.
- Gould, S. J., 1996. *The Mismeasure of Man*. Norton, New York.
- Gung, Y. C., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents. *Geophys. J. Int.*, **422**, 707–711.
- Gung, Y. C. & Romanowicz, B., 2004. Q tomography of the upper mantle using three component long period waveforms. *Geophys. J. Int.*, **157**, 813–830.
- Hadziioannou, C., Gaebler, P., Schreiber, K. U., Wassermann, J. & Igel, H., 2012. Examining ambient noise using colocated measurement of rotational and translational motion. *J. Seismol.*, **16**, doi: 10.1007/s10950-012-9288-5.
- Hartzell, H. S. & Heaton, T. H., 1983. Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Vally, California, earthquake. *Bull. Seismol. Soc. Am.*, **73**, 1553–1583.

- Hartzell, S., Liu, P., Mendoza, C., Ji, C. & Larson, K. M., 2007. Stability and Uncertainty of Finite-Fault Slip Inversions: Application to the 2004 Parkfield, California, Earthquake. *Bull. Seismol. Soc. Am.*, **97**, doi: 10.1785/0120070080.
- Hastings, W. K., 1970. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, **57**, 97–109.
- Hernandez, B., Cotton, F. & Campillo, M., 1999. Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake. *J. Geophys. Res.*, **104**, doi: 10.1029/1999JB900078.
- Hess, H. H., 1964. Seismic anisotropy of the uppermost mantle under oceans. *Nature*, **203**, 629–631.
- Ide, S., Beroza, G. C. & McGuire, J. J., 2005. Imaging earthquake source complexity. In A. Levander & G. Nolet (Eds.), *Seismic Earth: Array Analysis of Broadband Seismograms*, Volume 157, pp. 117–135. American Geophysical Union, John Wiley & Sons, Ltd.
- Igel, H., Cochard, A., Wassermann, J., Flaws, A., Schreiber, U., Velikoseltsev, A. & Dinh, N. P., 2007. Broad-band observations of earthquake-induced rotational ground motions. *Geophys. J. Int.*, **168**, doi: 10.1111/j.1365–246X.2006.03146.x.
- Igel, H., Nader, M., Kurrle, D., Ferreira, A., Wassermann, J. & Schreiber, K., 2011. Observations of Earth's toroidal free oscillations with a rotation sensor: The 2011 magnitude 9.0 Tohoku-Oki earthquake. *Geophys. Res. Lett.*, **38**, doi:10.1029/2011GL049045.
- Igel, H., Schreiber, U., Flaws, A., Schuberth, B., Velikoseltsev, A. & Cochard, A., 2005. Rotational motions induced by the M 8.1 Tokachi-oki earthquake, September 25, 2003. *Geophys. Res. Lett.*, **32**, doi:10.1029/2004GL022336.
- Ishii, M. & Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of Earth's mantle. *Science*, **285**, 1231–1236.
- Ishii, M. & Tromp, J., 2001. Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly. *Geophys. J. Int.*, **145**, 77–96.
- Iwata, T. & Sekiguchi, H., 2002. Source processes and near-source ground motion during the 2000 Tottori-ken Seibu earthquake. Presented at 11th Japan Earthquake Engineering Symposium, Earthquake Eng. Res. Liaison Comm., Sci Counc. of Jpn, Tokyo.
- Jackson, I., 2007. Physical origins of anelasticity and attenuation in rock. In G. Schubert (Ed.), *Treatise on Geophysics*, pp. 493–525. Elsevier.

- Jackson, I., Fitz Gerald, J. D., Faul, U. H. & Tan, B. H., 2002. Grain-size-sensitive seismic wave attenuation in polycrystalline olivine. *J. Geophys. Res.*, **107**, doi:10.1029/2001JB001225.
- Kendall, L. M., Langston, C. A., Lee, W. H. K., Lin, C.-J. & Liu, C.-C., 2012. Comparison of point and array-computed rotations for the TAIGER explosions of 4 March 2008. *J. Seismol.*, **16**, doi: 10.1007/s10950-012-9297-4.
- Kennett, B. L. N., 1997. Observational and theoretical constraints on crustal and upper mantle. *Phys. Earth. Planet. Int.*, **47**, 319–332.
- Kennett, B. L. N., 1998. On the density distribution within the Earth. *Geophys. J. Int.*, **132**, 374–382.
- Kennett, B. L. N. & Abdullah, A., 2011. Seismic wave attenuation beneath the Australasian region. *Austr. J. Earth Sci.*, **58**, 285–295.
- Kennett, B. L. N., Engdahl, E. R. & Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes. *Geophys. J. Int.*, **122**, 108–124.
- Kennett, B. L. N. & Sambridge, M. S., 1998. Inversion for multiple parameter classes. *Geophys. J. Int.*, **135**, 304–306.
- Kennett, B. L. N., Sambridge, M. S. & Williamson, P. R., 1988. Subspace methods for large inverse problems with multiple parameter classes. *Geophys. J. Int.*, **94**, 237–247.
- Khan, A., Zunino, A. & Deschamps, F., 2013. Upper mantle compositional variations and discontinuity topography imaged beneath australia from bayesian inversion of surface-wave phase velocities and thermochemical modeling. *J. Geophys. Res.*, **118**, doi:10.1002/jgrb.50304.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.*, **139**, 806–822.
- Komatitsch, D., Tsuboi, S. & Tromp, J., 2005. The spectral-element method in seismology. *Geophysical Monograph Series*, **157**, doi: 10.1029/156GM13.
- Kuo, C. & Romanowicz, B., 2002. On the resolution of density anomalies in the Earth's mantle using spectral fitting of normal mode data. *Geophys. J. Int.*, **150**, 162–179.
- Liu, C.-C., Huang, B., Lee, W. H. K. & Lin, C.-J., 2009. Observing rotational and translational ground motions at the hgsd station in taiwan from 2007 to 2008. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080156.
- Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods. *Geophys. J. Int.*, **174**, 265–286.

- Luo, Y. & Schuster, G. T., 1991. Wave-equation travelttime inversion. *Geophysics*, **56**, 645–653.
- Maceira, M. & Ammon, C. J., 2009. Joint inversion of surface wave velocity and gravity observations and its application to central asian basins shear velocity structure. *J. Geophys. Res.*, **114**, doi:10.29/2007JB005157.
- Mai, M., Burjanek, J., Delouis, B., Festa, G., Francois-Holden, C., Monelli, D., Uchide, T. & Zahradnik, J., 2007. Earthquake source inversion blindtest: Initial results and further developments. In *American Geophysical Union, Fall Meeting 2007, abstractS53C-08*. American Geophysical Union.
- Maurer, H., Curtis, A. & Boerner, D. E., 2010. Recent advances in optimized geophysical survey design. *Geophysics*, **75**, A177–A195.
- Mendoza, C. & Hartzell, S. H., 1989. Slip distribution of the 19 September 1985 Michoacan, Mexico, earthquake: Near-source and teleseismic constraints. *Bull. Seismol. Soc. Am.*, **79**, 655–669.
- Metropolis, N., Rosenbluth, M. N., Rosenbluth, A. W., Teller, A. H. & Teller, E., 1953. Equation of state calculations by fast computing machines. , **21**, 1087–1092.
- Metropolis, N. & Ulam, S., 1949. The Monte Carlo method. *J. Amer. Stat. Assoc.*, **44**, 335–341.
- Mikumo, T. & Aki, K., 1964. Determination of local phase velocity by intercomparison of seismograms from strain and pendulum instruments. *J. geophys. Res.*, **69**, 721–731.
- Monelli, D. & Mai, P. M., 2008. Bayesian inference of kinematic earthquake rupture parameters through fitting of strong motion data. *Geophys. J. Int.*, **173**, doi: 10.1111/j.1365–246X.2008.03733.x.
- Monelli, D., Mai, P. M., Jónsson, S. & Giardini, D., 2009. Bayesian imaging of the 2000 Western Tottori (Japan) earthquake through fitting of strong motion and GPS data. *Geophys. J. Int.*, **176**, doi: 10.1111/j.1365–246X.2008.03943.x.
- Montagner, J. P. & Anderson, D. L., 1989. Petrological constraints on seismic anisotropy. *Phys. Earth Planet. Int.*, **54**, 82–105.
- Mosca, I., Cobden, L., Deuss, A., Ritsema, J. & Trampert, J., 2012. Seismic and mineralogical structures of the lower mantle from probabilistic tomography. *J. Geophys. Res.*, **117**, doi:10.1029/2011JB008851.
- Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *J. Geophys. Res.*, **100**, doi: 10.1029/94JB03097.

- Nettles, M. & Dziewoński, A. M., 2008. Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America. *J. Geophys. Res.*, **113**, doi:10.1029/2006JB004819.
- Nigbor, R. L., Evans, J. R. & Hutt, C. R., 2009. Laboratory and field testing of commercial rotational seismometers. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080247.
- Nissen-Meyer, T., Fournier, A. & Dahlen, F. A., 2007. A two-dimensional spectral-element method for computing spherical-earth seismograms - I. Moment-tensor source. *Geophys. J. Int.*, **168**, 1067–1092.
- Olson, A. H. & Anderson, J., 1988. Implications of frequency-domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault. *Geophys. J. Int.*, **94**, 443–455.
- Olson, A. H. & Apsel, R. J., 1982. Finite faults and inverse theory with applications to the 1979 imperial valley earthquake. *Bull. Seismol. Soc. Am.*, **72**, 1969–2001.
- Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophys. J. Int.*, **167**, 361–379.
- Peyrat, S. & Olsen, K. B., 2004. Nonlinear dynamic rupture inversion of the 2000 Western Tottori, Japan, earthquake. *Geophys. Res. Lett.*, **31**, doi: 10.1029/2003GL019058.
- Pham, N. D., Huang, B.-S., Lin, C.-J., Vu, T.-M. & Tran, N.-A., 2012. Investigation of ground rotational motions caused by direct and scattered P-waves from the 4 March 2008 TAIGER explosion experiment. *J. Seismol.*, **16**, doi: 10.1007/s10950-012-9300-0.
- Pham, N. D., Igel, H., Wassermann, J., Käser, M., de la Puente, J. & Schreiber, U., 2009. Observations and modeling of rotational signals in the p coda: constraints on crustal scattering. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080101.
- Piatanesi, A., Cirella, A., Spudich, P. & Cocco, M., 2007. A global search inversion for earthquake kinematic rupture history: Application to the 2000 western Tottori, Japan earthquake. *J. Geophys. Res.*, **112**, doi: 10.1029/2006JB004821.
- Pillet, R., Deschamps, A., Legrand, D., Virieux, J., Béthoux, N. & Yates, B., 2009. Interpretation of broadband ocean-bottom seismometer horizontal data seismic background noise. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080123.
- Resovsky, J. & Ritzwoller, M., 1999. Regularisation uncertainty in density models estimated from normal mode data. *Geophys. Res. Lett.*, **26**, 2319–2322.
- Resovsky, J. & Trampert, J., 2003. Using probabilistic seismic tomography to test mantle velocity-density relationships. *Earth Planet. Sci. Lett.*, **215**, 121–134.

- Rickers, F., Fichtner, A. & Trampert, J., 2013. The Iceland - Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion. *Earth Planet. Sci. Lett.*, in press.
- Ritzwoller, M. H., 2009. Ambient noise seismic imaging. *McGraw Hill Yearbook of Science and Technology*.
- Romanowicz, B., 2001. Can we resolve 3D density heterogeneity in the lower mantle? *Geophys. Res. Lett.*, **28**, 1107–1110.
- Ruan, Y. & Zhou, Y., 2010. The effects of 3-D anelasticity (Q) structure on surface wave phase delays. *Geophys. J. Int.*, **181**, 479–492.
- Sacks, I. S., Snoke, J. A., Evans, R., King, G. & Beavan, J., 1976. Single-site phase velocity measurement. *Geophys. J. R. astr. Soc.*, **46**, doi: 10.1111/j.1365–246X.1976.tb04157.x.
- Sambridge, M. S. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems. *Rev. Geophys.*, **40**, doi: 10.1029/2000RG000089.
- Schreiber, U., Hautmann, J. N., Velikoseltsev, A., Wassermann, J., Igel, H., Otero, J., Vernon, F. & Wells, J.-P. R., 2009. Ring laser measurements of ground rotations for seismology. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080171.
- Schreiber, U., Stedman, G., Igel, H. & Flaws, A., 2006. Ring laser gyroscopes as rotation sensors for seismic wave studies. In R. Teisseyre, M. Takeo, & E. Majewski (Eds.), *Earthquake Source Assymetrie, Structural Media and Rotation Effects*, pp. 377–390. Springer, Heidelberg.
- Schreiber, U., Velikoseltsev, A., Carr, A. J. & Franco-Anaya, R., 2009. The application of fiber optic gyroscopes for the measurement of rotations in structural engineering. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080086.
- Semmane, F., Cotton, F. & Campillo, M., 2005. The 2000 Tottori earthquake: A shallow earthquake with no surface rupture and slip properties controlled by depth. *J. Geophys. Res.*, **110**, doi: 10.1029/2004JB003194.
- Shapiro, N. M., Campillo, M., Stehly, L. & Ritzwoller, M., 2005. High resolution surface wave tomography from ambient seismic noise. *Science*, **307**, 1615–1618.
- Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007a. Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods. *Geophys. J. Int.*, **171**, 368–389.
- Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007b. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods. *Geophys. J. Int.*, **168**, 1153–1174.

- Sieminski, A., Trampert, J. & Tromp, J., 2009. Principal component analysis of anisotropic finite-frequency kernels. *Geophys. J. Int.*, **179**, 1186–1198.
- Simmons, N. A., Forte, A. M., Boschi, L. & Grand, S. P., 2010. GyPSuM: A joint tomography model of mantle density and seismic wave speeds. *J. Geophys. Res.*, **115**, doi:10.1029/2010JB007631.
- Somerville, P. G., Smith, N. F., Graves, R. W. & Abrahamson, N. A., 1997. Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. *Seis. Res. Lett.*, **68**, 199–222.
- Spudich, P. & Archuleta, R. J., 1987. Techniques for earthquake ground motion calculation with application to source parameterization of finite faults. In B. Bolt (Ed.), *Seismic Strong Motion Synthetics, Computational Techniques*, pp. 205–265. Academic Press, New York.
- Spudich, P. & Fletcher, J. B., 2008. Observation and Prediction of Dynamic Ground Strains, Tilts, and Torsions Caused by the Mw 6.0 2004 Parkfield, California, Earthquake and Aftershocks, Derived from UPSAR Array Observations. *Bull. Seismol. Soc. Am.*, **98**, doi: 10.1785/0120070157.
- Spudich, P., Steck, L. K., Hellweg, M., Fletcher, J. B. & Baker, L. M., 1995. Transient stresses at Parkfield, California, produced by the M 7.4 Landers earthquake of June 28, 1992: Observations from the UPSAR dense seismograph array, *J. Geophys. Res.*, **100**(b1), 675–690, *J. Geophys. Res.*, **100**, doi: 10.1029/94JB02477.
- Stein, S., Geller, R. & Liu, M., 2012. Why earthquake hazard maps often fail and what to do about it. *Tectonophysics*, doi: 10.1016/j.tecto.2012.06.047.
- Stupazzini, M., de la Puente, J., Smerzini, C., Käser, M., Igel, H. & Castellani, A., 2009. Study of rotational ground motion in the near-field region. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080153.
- Suryanto, W., Igel, H., Wassermann, J., Cochard, A., Schuberth, B., Vollmer, D., Scherbaum, F., Schreiber, U. & Velikoseltsev, A., 2006. First comparison of array-derived rotational ground motions with direct ring laser measurements. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120060004.
- Takeo, M., 1998. Ground rotational motions recorded in near-source region of earthquakes. *Geophys. Res. Lett.*, **25**, doi: 10.1029/98GL00511.
- Takeo, M., 2009. Rotational motions observed during an earthquake swarm in april 1998 offshore Ito, Japan. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080173.
- Takeuchi, H. & Saito, M., 1972. Seismic surface waves. in: *Methods in Computational Physics*, editor B. A. Bolt, **11**, 217–295.

- Taleb, N. N., 2007. *The Black Swan: The Impact of the Highly Improbable*. Random House, New York.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust. *Science*, **325**, 988–992.
- Tarantola, A., 1986. A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, **51**, 1893–1903.
- Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *Pure Appl. Geophys.*, **128**, 365–399.
- Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation, 2nd edition*. Society for Industrial and Applied Mathematics, Philadelphia.
- Tarantola, A., 2006. Popper, Bayes and the inverse problem. *Nature Physics*, doi:10.1038/nphys375.
- Tinti, E., Cocco, M., Fukuyama, E. & Piatanesi, A., 2009. Dependence of slip weakening distance (D_c) on final slip during dynamic rupture of earthquakes. *Geophys. J. Int.*, **177**, 1205–1220.
- Tondi, R., Achauer, U., Landes, M., Davi, R. & Besutiu, L., 2009. Unveiling seismic and density structure beneath the Vrancea seismogenic zone, Romania. *J. Geophys. Res.*, **141**, doi:10.1029/2008JB005992.
- Tondi, R., de Franco, R. & Barzaghi, R., 2000. Sequential inversion of refraction and wide-angle reflection traveltimes and gravity data for two-dimensional velocity structures. *Geophys. J. Int.*, **141**, 679–698.
- Trampert, J., Deschamps, F., Resovsky, J. & Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle. *Science*, **306**, 853–856.
- Trampert, J. & Fichtner, A., 2013. Global imaging of the Earth's deep interior: seismic constraints on (an)isotropy, density and attenuation. In S. Karato (Ed.), *Physics and Chemistry of the deep Earth*. Wiley-Blackwell.
- Trampert, J. & van Heijst, H. J., 2002. Global azimuthal anisotropy in the transition zone. *Science*, **296**, 1297–1299.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.*, **160**, 195–216.
- Velikoseltsev, A., Schreiber, K. U., Yankovsky, A., Wells, J.-P. R., Boronachin, A. & Tkachenko, A., 2012. On the application of fiber optic gyroscopes for detection of seismic rotations. *J. Seismol.*, **16**, doi: 10.1007/s10950-012-9282-y.

BIBLIOGRAPHY

- Wald, D. J. & Heaton, T. H., 1994. Spatial and temporal distribution of slip for the 1994 Landers, California earthquake. *Bull. Seismol. Soc. Am.*, **84**, 668–691.
- Wang, H., Igel, H., Gallovic, F. & Cochard, A., 2009. Source and basin effects on rotational ground motions: Comparison with translations. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/01200801.
- Wassermann, J., Lehndorfer, S., Igel, H. & Schreiber, U., 2009. Performance test of a commercial rotational motions sensor. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080157.
- Wu, C.-F., Lee, W. H. K. & Huang, H. C., 2009. Array deployment to observe rotational and translational ground motions along the meishan fault, taiwan: a progress report. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080185.
- Wu, R. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, **50**, 582–595.
- Yoshizawa, K. & Kennett, B. L. N., 2004. Multi-mode surface wave tomography for the Australian region using a 3-stage approach incorporating finite-frequency effects. *J. Geophys. Res.*, **109**, doi:10.1029/2002JB002254.
- Zembaty, Z., 2009. Rotational Seismic Load Definition in Eurocode 8, Part 6 for Slender Tower-shaped Structures. *Bull. Seismol. Soc. Am.*, **99**, doi: 10.1785/0120080252.
- Zhou, Y., 2009. Surface-wave sensitivity to 3-D anelasticity. *Geophys. J. Int.*, **178**, 1403–1410.
- Zhou, Y., Dahlen, F. A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables. *Geophys. J. Int.*, **158**, 142–168.
- Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2012. Structure of the European upper mantle revealed by adjoint tomography. *Nat. Geosc.*, **5**, 493–498.

Acknowledgments

"Science, since people must do it, is a socially embedded activity. It progresses by hunch, vision, and intuition. Much of its change through time does not record a closer approach to absolute truth, but the alteration of cultural contexts that influence it so strongly."
(Gould, 1996)

In 1859 Bernhard Riemann proposed the hypothesis that the nontrivial zeros of the Riemann zeta function all have real part $1/2$ - a conjecture that has thus far resisted all attempts to prove or disprove it. Understanding such unsolved problems at the edge of knowledge marked the fascinating highlights for me as an undergraduate student of mathematics. For a Ph.D. student unsolved problems are omnipresent and sometimes more frustrating than fun. In any case I very much appreciate that my advisors Heiner Igel and Andreas Fichtner offered me a unique environment for exploring this universe of challenges. Thank you, Heiner, your enthusiasm and energy for scientific questions and beyond is admirable. Thank you, Andreas, for guiding me with infinite patience and sharing your talents. That was not a given.

This work also benefited substantially from Jeannot Trampert's assistance, who accepted me as a member of his seismology group at Utrecht University for half a year. Andrew, Denise, Florian, Paul and Ralph, meeting you there and at numerous workshops and conferences was a gift.

It was a pleasure to be part of the institute in Munich. The diversity of my colleagues and office mates was life-enhancing. Lion, you finally convinced me that computers in principle do have the potential to simplify life. Mafe and Lorenzo, I appreciate not only your bravery and open mind for listening with me to *Kofelgschroa*. For providing an extremely useful data processing toolbox my thanks also go to the ObsPy people.

Of course, such a project is not conceivable without professional technical support. Thank you Alex, Gerald and Jens. The IT-infrastructure at our institute is outstanding. However, concerning the supernatural computing tasks I render homage to the creators of SuperMUC.

Sorry, Greta, Helen and Yvonne for all the paperwork I produced. I really acknowledge your diligent help.

Finally, I would like to give special thanks to my cozy alien. Waking me up every day at 5 a.m. substantially pushed this project.